

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N	A2	(11) International Publication Number: WO 97/44439 (43) International Publication Date: 27 November 1997 (27.11.97)
--	-----------	---

(21) International Application Number: PCT/US97/08690

(22) International Filing Date: 22 May 1997 (22.05.97)

(30) Priority Data:

08/653,650

24 May 1996 (24.05.96)

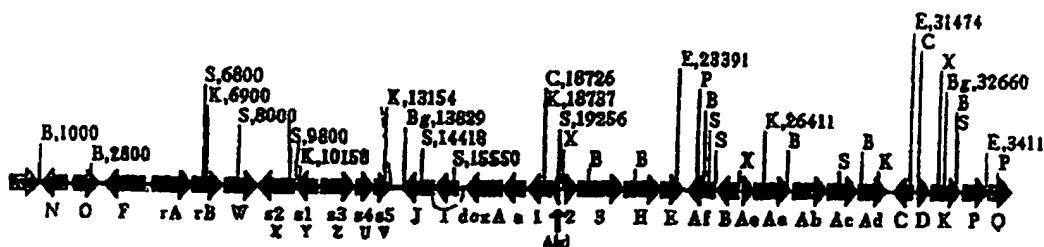
US

(71) Applicant: THE OHIO STATE UNIVERSITY RESEARCH
FOUNDATION [US/US]; 1960 Kenny Road, Columbus,
OH 43210-1063 (US).(72) Inventors: STROHL, William, A.; 8462 Invergordon Court,
Dublin, OH 43017 (US). DICKENS, Michael, L.; 1228
Wimbeldon Boulevard, Columbus, OH 43228 (US). De-
SANTI, Charles, L.; 2276 Northwest Boulevard, Columbus,
OH 43212 (US).(74) Agent: GOLRICK, Mary, E.; Calfee, Halter & Griswold LLP,
Suite 1400, 800 Superior Avenue, Cleveland, OH 44114
(US).(81) Designated States: AU, CA, JP, European patent (AT, BE,
CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL,
PT, SE).

Published

Without international search report and to be republished
upon receipt of that report.

(54) Title: METHOD OF PRODUCING DOXORUBICIN



(57) Abstract

The present invention provides novel methods for producing doxorubicin using daunomycin as a substrate. One method employs a genetically engineered host microorganism which is transformed with a vector, preferably a plasmid, which contains the *doxA* gene. Preferably, the *doxA* gene, also referred to herein as "*doxA*", is cloned into a plasmid which is then introduced into the host microorganism, preferably a bacterial host, more preferably *Streptomyces*, to provide a transformed host microorganism. The *doxA* gene, when present on a plasmid, confers on the transformed host the ability to convert daunomycin and 13-dihydrodaunomycin, to doxorubicin. The *doxA* gene encodes a P450-like enzyme which catalyzes the hydroxylation of daunomycin and 13-dihydrodaunomycin at C-14 to form doxorubicin; such enzyme is designated "daunomycin C-14 hydroxylase". Thus, the expression of *doxA* in the transformed host using a plasmid which contains *doxA* enables the transformed host to convert daunomycin to doxorubicin. The doxorubicin is then extracted from host microorganism cultures.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

METHODS OF PRODUCING DOXORUBICIN

Background of the Invention

Daunomycin and doxorubicin are clinically important chemotherapeutic agents. Daunomycin is used primarily to treat adult myelogenous leukemia. Doxorubicin is widely used to treat a variety of neoplasias, making it the more valuable of the two anticancer drugs.
5 The world wide market for doxorubicin is estimated to exceed \$156 million. As of 1984, the wholesale price for doxorubicin was estimated to be \$1,370,000 per kilogram.

While daunomycin is synthesized by several species of *Streptomyces*, doxorubicin is biologically synthesized by only one
10 strain, a mutant strain of *Streptomyces peucetius*, called *S. peucetius subsp. caesius* which is available from the American Type Culture Collection under Accession number 27952.

The alternative in vitro laboratory synthesis of doxorubicin is difficult. The in vitro synthesis of doxorubicin is a process
15 involving multiple steps and resulting in a poor yield, with a lack of stereospecificity in several of the synthetic steps, producing forms which are difficult to separate.

Chemical synthetic procedures are known for converting daunomycin to doxorubicin; however they require the use of halogens in
20 the synthetic process.

It would be desirable to have an efficient, cost-effective method for producing doxorubicin that does not require the use of halogens in the synthetic process.

Summary of the Invention

25 The present invention provides novel methods for producing doxorubicin using daunomycin as a substrate. One method employs a genetically engineered host microorganism which is transformed with a vector, preferably a plasmid, which contains the *doxA* gene. Preferably, the *doxA* gene, also referred to herein as "*doxA*", is
30 cloned into a plasmid which is then introduced into the host microorganism, preferably a bacterial host, more preferably *Streptomyces*, to provide a transformed host microorganism. The *doxA* gene, when present on a plasmid, confers on the transformed host the ability to convert daunomycin and 13-dihydrodaunomycin, to
35 doxorubicin. The *doxA* gene encodes a cytochrome P450-type enzyme

-2-

which catalyzes the hydroxylation of daunomycin and 13-dihydrodaunomycin at C-14 to form doxorubicin; such enzyme is designated "daunomycin C-14 hydroxylase". Thus, the expression of *doxA* in the transformed host using a plasmid which contains *doxA* enables the transformed host to convert daunomycin to doxorubicin. The doxorubicin is then extracted from host microorganism cultures.

Another method for producing doxorubicin involves incubating the daunomycin C-14 hydroxylase with daunomycin, then extracting the doxorubicin from the solution.

Another method involves adding daunomycin to cultures of *Streptomyces* sp. strain C5 and extracting doxorubicin from the culture fluid and the host cells.

The invention also relates to daunomycin C-14 hydroxylase, novel plasmids, novel polylinkers and novel transformed host microorganisms employed in such method for producing doxorubicin. The invention also relates to methods for producing anthracyclines, such as 13-deoxycarminomycin and 13-deoxydaunomycin, 13-dihydrocarminomycin and 13-dihydrodaunomycin, carminomycin and daunomycin.

Brief Description of Figures

Figure 1 is a restriction map of the *Streptomyces* sp. strain C5 daunomycin biosynthesis gene cluster which shows the position of *doxA* within the cluster. Abbreviations for restriction endonuclease sites are as follows: "B" represents *Bam*HI; "Bg" represents *Bgl*II; "C" represents *Cla*I; "E" represents *Eco*RI; "K" represents *Kpn*I; "P" represents *Pst*I; "S" represents *Sst*I; and "X" represents *Xho*I;

Figure 2 is a detailed restriction map of part of the daunomycin biosynthesis gene cluster from *Streptomyces* sp. strain C5. Abbreviations for restriction endonuclease sites are as follows: "B" represents *Bam*HI; "Bg" represents *Bgl*II; "E" represents *Eco*RI; "K" represents *Kpn*I; "P" represents *Pst*I; "S" represents *Sst*I; "Sp" represents *Sph*I;

Figure 3 is a nucleotide sequence of the 3196 base pair *Kpn*I-*Sst*I DNA fragment from *Streptomyces* sp. strain C5 containing the *doxA* gene. The deduced amino acid sequence of the daunomycin C-14 hydroxylase is given below the nucleotide sequence. Potential ribosome binding sites, designated "rbs" are identified, as are significant restriction endonuclease sites. The sequences and deduced products of the 3' end of *orf1*, all of *orfA*, and the 5' end of *dauI* are also shown;

Figure 4 shows the plasmid maps of plasmid pANT849 and the plasmids pANT42 and pANT842 which were used to construct pANT849;

Figure 5 shows the plasmid maps of pANT195 and plasmids pANT849, pANT186, pANT185, pANT235 and pUC19 all of which were used to construct pANT195;

Figure 6 shows the sequence of *snpR*, *doxA*, and the intervening sequences within plasmid pANT195;

Figure 7 shows plasmid maps of pANT192 and pANT193;

Figure 8 shows plasmid maps of pANT194 and pANT196;

Figure 9 shows the N-terminal, modified region of the *doxA* fusion protein;

Figure 10 shows plasmid maps of pANT198 and pANT199; and

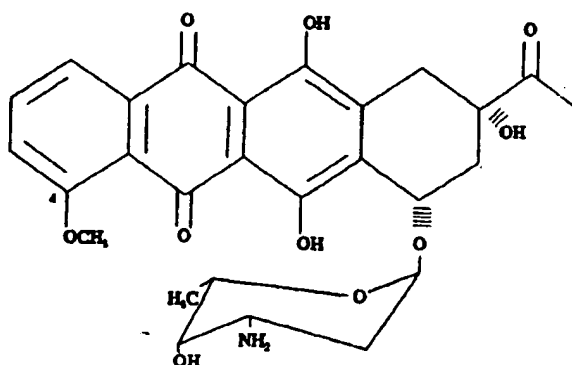
Figure 11 shows sequence of the *doxA* gene and upstream *melC1* promoter region in pANT196; and

Figure 12 shows a plasmid map of pANT144.

Detailed Description of Invention

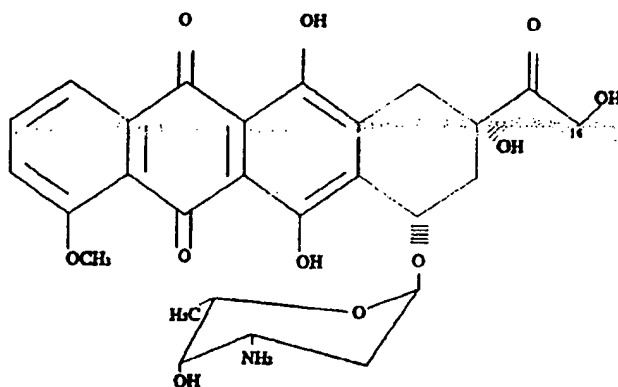
A novel method for producing doxorubicin from daunomycin has been developed which employs genetically engineered host microorganisms that contain and express the gene *doxA*. Preferably, *doxA* is cloned into a plasmid which is then inserted into a host microorganism, preferably a bacterial host, more preferably *Streptomyces*, most preferably *Streptomyces lividans* to provide a transformed host. The *doxA* gene encodes daunomycin C-14 hydroxylase which catalyzes the hydroxylation of daunomycin and 13-dihydrodaunomycin at C-14 to form doxorubicin. Thus, expression of *doxA* in the transformed host using a plasmid which contains the *doxA* gene enables the transformed host to convert daunomycin to doxorubicin.

Daunomycin is also known as daunorubicin; doxorubicin is also known as 14-hydroxydaunomycin and adriamycin. The structure of daunomycin is shown below:



-4-

The structure of doxorubicin is shown below:



Cloning and Analysis of the *doxA* Gene

Streptomyces sp. strain C5 synthesizes several compounds in fermentations, that is such compounds are produced from common metabolic intermediates and without the addition of precursor anthracycline molecules to the culture media. *Streptomyces* sp. strain C5 produces the following anthracyclines: ϵ -rhodomycinone; daunomycin; 13-dihydrodaunomycin; baumycin A1; and baumycin A2. Nevertheless, a gene, the *doxA* gene, was discovered in the genome of *Streptomyces* sp. strain C5 which, when expressed, converts daunomycin, particularly exogenous daunomycin, to doxorubicin. Preferably, the conversion of daunomycin to doxorubicin is accomplished by cloning the *doxA* gene along with a promoter into a plasmid which is then introduced into a host microorganism.

-5-

It has also been discovered that *Streptomyces* sp. strain C5 can convert small amounts, less than 10%, daunomycin to doxorubicin in the absence of plasmid containing the *doxA* gene.

Preferably, the *doxA* gene is cloned from *Streptomyces*, preferably *Streptomyces* sp. strain C5. Alternatively, the *doxA* gene is synthesized using conventional oligonucleotide synthesis techniques and equipment.

The *doxA* gene is located in the daunomycin biosynthesis gene cluster between the daunomycin polyketide biosynthesis genes and *dauI*, a putative transcriptional activator as shown in Figure 1. The location of *doxA* within the *Streptomyces* sp. strain C5 daunomycin biosynthesis gene cluster is shown in Figure 2.

The approximately 8 kbp region between *dauI*, a gene encoding an activator regulatory protein for daunomycin biosynthesis, and the daunomycin polyketide synthase biosynthesis genes was sequenced in its entirety.

Plasmids containing inserts to be sequenced were isolated from recombinant *E. coli* JM83, available from Dr. Mary Berlyn, *E. coli* Genetic Stock Center, Yale University, P.O. Box 6666, New Haven, Ct 06511-7444 by the methods disclosed in Carter, M.J., and I.D. Milton, (1993), "An Inexpensive and Simple Method for DNA Purification on Silica Particles," *Nucleic Acids Res.* Volume 21, p. 1044. The *doxA* DNA was sequenced in both directions, that is, both strands were sequenced using Sequenase enzyme, Version 2.0 from the United States Biochemical Corp., Cleveland, Ohio, according to the manufacturer's instructions, and as described in Ye, et. al., 1994, "Isolation and Sequence Analysis of Polyketide Synthase Genes from the Daunomycin-producing *Streptomyces* sp. strain C5" *J. Bacteriol.* 176:6270-6280. Doubled-stranded DNA templates were employed. The terminated chains were labeled with 3000 Ci/mmol (α -³²P)dCTP from Dupont-New England Nuclear, Boston, MA. The terminated labeled chains were separated on a 6% weight-to-volume polyacrylamide gel containing 10% (volume-to-volume) formamide and visualized by autoradiography. Sequencing reactions were carried out using 7-deaza-dGTP nucleotide mixes to reduce compressions. Forward (-40) and reverse universal pUC/m13 17-mer oligonucleotide primers from U.S. Biochemical Corp. were used to obtain the initial sequences in the inserts. Specific primers, 15-mer oligonucleotides, were generated based on sequencing results for extension of the sequences within the inserts.

DNA sequence data were analyzed using Clone Manager from Stateline, PA, and the Sequence Analysis Software Package of the Genetics Computer Group from Madison, WI.

The nucleotide sequence between *dauI* and the ketoreductase just downstream of *dauA-orfG* is shown in Figure 3. Two complete open

-6-

reading frames, *orfA* and *doxA*, were found within this sequence; *OrfA* encodes a protein of M_r 28,808, and 275 amino acid residues, and *doxA* encodes a protein of M_r 46,096 and 422 amino acid residues.

Plasmids

5 The *doxA* gene is inserted into a vector, preferably a plasmid. Optionally, the plasmid contains genes from the daunomycin synthesis cluster in addition to *doxA*. However, preferred plasmids lack *dauA(g)* and more preferred plasmids lack *dauA(g)*, *orf1* and *orfA*.

10 The preferred plasmids contain not only the translated portion of *doxA* but a promoter. Suitable promoters include, *Streptomyces* promoters for example, *melC1-P*, *ermE-P*, wild type, and *snpA-P*. The *snpA-P* promoter is the most preferred. Preferably, the promoter is a protein activated promoter, and most preferably, an SnpR-activated promoter. Less preferred plasmids, such as pANT196, contain a *melC1*
15 promoter from pIJ702 for expression of *doxA*. Also less preferred are plasmids which lack a known promoter, such as pANT194.

 The most preferred plasmid which contains *doxA* is designated "pANT195" which is shown in Figure 5. Host microorganisms, when transformed with plasmid pANT195, convert 100% daunomycin to
20 doxorubicin. Other plasmids which contain *doxA* are suitable, including, for example pANT192, pANT193, pANT194 and pANT196. Host microorganisms, when transformed with plasmid pANT192, typically convert about 25% daunomycin to doxorubicin at a concentration of 2 μ g/ml. Host microorganisms when transformed with pANT193 convert about
25 80% daunomycin to doxorubicin and about 20% daunomycin to 13-dihydrodaunomycin, at a daunomycin concentration of 2 μ g/ml.

Construction of the Plasmids

 Digestion of and ligation of DNA was performed using conventional techniques described by Maniatis et al. (1982) in
30 "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory, N.Y.

Construction of pANT195

 Plasmid pANT195, shown in Figure 5, has about 7.04 kbp of DNA. Plasmid pANT195 was constructed by inserting the 1.72 kbp *SphI*-*SacI*
35 fragment insert containing intact *doxA* from plasmid pANT186 into pANT849.

 First, plasmid pANT186 was constructed by constructing pANT235. Plasmid pANT235 is described in Ye et. al. 1994 "Isolation and Sequence Analysis of Polyketide Synthase Genes from the Daunomycin-Producing
40 *Streptomyces* sp. Strain C5" *J. Bacteriol.* Vol. 176, pp. 6270-6280. Plasmid pANT235 is a 9.2 kbp plasmid which contains a 6.48 kbp *BamHI*-*BglII* DNA fragment from the *Streptomyces* sp. strain C5 daunomycin biosynthesis gene cluster. The *doxA* gene lies within the insert of pANT235 which is derived from the daunomycin biosynthesis gene cluster.

-7-

The *Bam*HI-*Bgl*III DNA fragment had been cloned into the *Bam*HI site of pUC19 to generate pANT235. Plasmid pUC19 is available from Gibco BRL, Gaithersburg MD.

Next, pANT235 was digested with *Sal*I and *Sst*I and the digestion products were purified on an agarose gel. The 1.67 kbp *Sal*I-*Sst*I fragment containing the 3' end of the *doxA* gene and the 5' end of *dauI* was extracted from the agarose gel and ligated into pUC19 with T4 DNA ligase, from Gibco BRL, Gaithersburg MD, to generate pANT185, as shown in Figure 5.

Next, pANT235 was used as the template for the polymerase chain reaction amplification of the 5'-end of the *doxA* gene containing an upstream ribosome binding site and *Sph*I restriction site for the 5' end and *Bsp*EI restriction site for the 3' end.

The forward primer used in the polymerase chain reaction amplification of the *doxA* gene had the following nucleotide sequence: 5'-GACATGCATGCGGAGGGTGCCTC-3' SEQ.ID 1

The forward primer which is used for the 5'-end, contains an *Sph*I site with five extra nucleotides on the end and the extra ribosome binding site "GGAGG". The reverse primer had the following nucleotide sequence:

5'-GACGCAGCTCCGGAACGGGG-3' SEQ.ID 2

The reverse primer which is used for the 3'-end, has a *Bsp*EI site plus eight extra nucleotides.

The polymerase chain reaction amplification was carried out for 25 cycles using Deep Vent Polymerase from New England Biolabs, Beverly, MA. The solution for PCR included: 2.0 μ l dimethylsulfoxide; 14.5 μ l double distilled water; 1.25mM dNTPs, 16.0 μ l of a total stock containing dATP, dCTP, dTTP, dGTP; 5.0 μ l 10X Deep Vent Buffer, from New England Biolabs; 5 μ l forward primer; 5 μ l reverse primer; 0.5 μ l Deep Vent polymerase; 2.0 μ l DNA template, 14.5 μ l distilled water, and boiled for 10 minutes. PCR was carried out by incubating the reaction mixture at 94°C in the absence of Deep Vent Polymerase for 5 minutes, following by 25 cycles of the following regimen: 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds. After 25 cycles were completed, the mixture was incubated at 72°C for 7 minutes and then held at 4°C until further use.

The products of PCR were separated on a 0.8% agarose gel, and a 298 base pair DNA fragment was eluted from the gel. The 298 base pair DNA fragment was then digested with *Sph*I and *Bsp*EI to generate a 285 base pair fragment with "sticky" ends. pANT185 was digested with *Sph*I and *Bsp*EI and the 285 base pair fragment ligated into pANT185 to generate pANT186 which was introduced into *dam/dcm*-minus *E. coli* strain ET12567. MacNeil, et. al. (1992) "Analysis of *Streptomyces avermitilis* Genes Required for Avermectin Biosynthesis Utilizing a

-8-

Novel Integration Vector", *Gene*, vol. 111, pages 61-68. Plasmid pANT186 contains the complete *doxA* gene, upstream of which lay the newly constructed ribosome binding site having the nucleotide sequence GGAGG. The nucleotide sequence of the PCR-generated 5' end of the gene was confirmed by dideoxy sequencing.

Plasmid pANT849, shown in Figure 4, was constructed by first constructing pANT842. Plasmid pANT42, described in Lampel, et. al., 1992 "Cloning and Sequencing of a Gene Encoding a Novel Extracellular Neutral Proteinase from *Streptomyces* sp. Strain C5 and Expression of the Gene in *Streptomyces lividans* 1326" *J. Bacteriology* 174:2797-2808, was digested with *KpnI* and religated, removing a 1.95 kbp *KpnI* fragment to yield pANT842.

A novel polylinker sequence, having 48 nucleotides, was constructed according to conventional techniques using synthesized DNA oligonucleotides by Integrated DNA Technologies, Inc., Coralville, Iowa. The polylinker sequence has the following nucleotide sequence:

SphI *BglII* *SacI* *DraI* *HpaI*
GCATGCGAATTCAGATCTAGAGCTCAAGCTTTAACTAGTTAACGCGT SEQ.ID 3
EcoRI XbaI HindIII SpeI MluI

Plasmid pANT842 was digested with *SphI*-*MluI* to remove a 1.42 kbp *SphI*-*MluI* fragment. The polylinker sequence was ligated into *SphI*-*MluI*-digested pANT842 to provide plasmid pANT849. Plasmid pANT849, shown in Figure 4, has 5.34 kbp of DNA and lacks the *snpA* gene and most of *melC2*. Plasmid pANT849 does have the *Snpr*-activated *snpA*-promoter, which is located immediately upstream of the polylinker sequence as shown in Figure 4. pANT849 is a high copy number plasmid and contains the thiostrepton resistance gene as the selectable marker.

Next, to construct pANT195, a clone of pANT186 which contains the modified *doxA* gene, was digested with *SphI* and *SstI*, and pANT849 was digested with *SphI* and *SstI*. The fragment from pANT186 containing the *doxA* gene was ligated into the polylinker sequence of pANT849 to make pANT195 as shown in Figure 5.

The sequence of the region of pANT195 containing the *snpr* activator gene, the *Snpr*-activated *snpA* promoter, and the 5'-end-modified *doxA* gene is shown in Figure 6, and Seq. Id. No. 6.

Construction of pANT192

Plasmid pANT192 shown in Figure 7 is an 11.84 kbp plasmid which contains DNA encoding the acyl carrier protein and its putative promoter, a ketoreductase (*orf1*), *orf2*, a partial *orf3*, *orfA*, *doxA*, *dauI*, and most of *dauJ*. Plasmid pANT192 was constructed by removing the 6.52 kbp *HindIII*-*EcoRI* fragment from pANT235 which includes the entire *BglIII*-*BamHI* fragment, by digesting pANT235 with *HindIII* and *EcoRI*. Next pANT849 was digested with *HindIII* and *EcoRI* and the 6.52 kbp *HindIII*-*EcoRI* fragment from pANT235 was ligated into pANT849.

-9-

Construction of pANT193

Plasmid pANT193, shown in Figure 7, has 10.28 kbp of DNA and contains part of *orf1*, all of *orfA*, and *doxA* driven by the *snpR*-activated *snpA*-promoter. Plasmid pANT193 was constructed by digesting
5 plasmid pANT235 with *KpnI*, and the 1582 base pair *KpnI* fragment removed. The plasmid was re-ligated to itself to form pANT235-k. Plasmid pANT235-k was digested with *EcoRI* and *HindIII* to remove the 4.95 kbp *EcoRI*-*HindIII* fragment. pANT849 was digested with *EcoRI* and *HindIII* and the 4.95 kbp *EcoRI*-*HindIII* fragment from pANT235-k was
10 ligated into the digested pANT849.

Construction of pANT194

Plasmid pANT194, shown in Figure 8, has 8.97 kbp of DNA and contains the part of *orf1*, all of *orfA*, *dauI*, *dauJ* and *doxA* but lacks any known promoter to drive the expression of *doxA*. Plasmid pANT194
15 was constructed by digesting pANT192 with *KpnI* to remove a 2.87 kbp *KpnI* fragment and then religating the plasmid to itself.

Construction of pANT196

Plasmid pANT196, shown in Figure 8, has 7398 bp of DNA and possesses a promoter *melC1* which drives the expression of the *doxA*
20 gene. pANT186 was digested with *SphI* and *SstI* and a 1712 nucleotide *SphI*-*SstI* fragment from pANT186 containing *doxA* was isolated and ligated into *SphI*-*SstI* digested pIJ702. Plasmid pIJ702 is a 5.686 kbp plasmid which is described in Katz, E., et. al. (1983) "Cloning and Expression of the Tyrosinase Gene from *Streptomyces antibioticus* in
25 *Streptomyces lividans*" J. Gen. Microbiol. volume 129, pages 2703-2714.

Construction of pANT198

Plasmid pANT186 was digested with *SphI* and then incubated with T4 DNA polymerase from Gibco BRL, according to the manufacturer's
30 instructions to yield a blunt end. Plasmid pZero from Invitrogen, San Diego, Ca., was digested with *EcoRI* and then filled in 5' to 3' using Klenow fragment of DNA polymerase according to the manufacturer's instructions to provide a blunt end. Both fragments were purified according to the methods described in Carter, M.J. and I.D. Milton
35 (1993), Nucleic Acids Res. volume 21, pages 1044. The fragments were precipitated in ethanol for one hour at -70°C and then digested with *SstI* overnight. The plasmid and insert, each of which contains a single blunt end and an *SstI* end, were purified from an agarose gel and then ligated overnight with T4 DNA ligase at room temperature, to
40 provide pANT198.

Construction of pANT199

Plasmid pANT198 was digested with *EcoRI*-*HindIII*, the *EcoRI*-*HindIII* fragment removed and ligated into pTrcHisC from Invitrogen to construct pANT199. In pANT199, the *doxA* gene is translationally fused

-10-

with a leader sequence encoding six histidine residues, shown in SEQ. ID. NO. 7 so that the fusion protein can be affinity purified on a nickel-agarose gel.

pANT849

- 5 Plasmid pANT849 in addition to being useful to construct pANT195 is also useful expression vector for other genes. To construct other such plasmids pANT849 is digested with at least one restriction endonuclease corresponding to the restriction sites in the polylinker, such as, for example, *SphI*, *BglII*, *SacI*, *DraI*, *HpaI*, *EcoRI*, *XbaI*,
10 *HindIII*, *SpeI*, or *MluI*. The desired gene sequence to be inserted into the plasmid is provided with sticky ends corresponding to the sticky ends of the cut pANT849. The desired gene is then ligated into the plasmid to provide a new plasmid derived from pANT849.

Host Microorganisms

- 15 Suitable host microorganisms for the *doxA* plasmid possess electron donating, cytochrome P450 accessory proteins; suitable accessory proteins include for example, NADPH:ferredoxin oxidoreductase and ferredoxin. The preferred host microorganisms are bacteria, more preferably *E. coli* or *Streptomyces* spp., most preferably *Streptomyces*
20 *lividans* TK24 and *Streptomyces coelicolor* CH999. *Streptomyces coelicolor* CH999 is a mutant of *Streptomyces coelicolor* A3(2). *Streptomyces lividans* TK4 is available from Professor David A. Hopwood, Head, Department of Genetics, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, United Kingdom. *Streptomyces coelicolor* CH999
25 is available from C. Khosla, Stanford University, Palo Alto, California and Professor David A. Hopwood. *S. peucetius* is available from the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland, USA under the accession number 29050. *Streptomyces* sp. strain C5 was obtained from the Frederick Cancer Research Center,
30 Frederick MD. *Streptomyces* sp. strain C5 is also available from the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland, USA under the Accession number ATCC 49111.

Transformation of Host Microorganism

- The plasmids are introduced into the host microorganism using
35 conventional techniques. For example, *Streptomyces* spp. are transformed using electroporation as described in Pigac and Schrempf (1995) "A Simple and Rapid Method of Transformation of *Streptomyces rimosus* R6 and Other Streptomycetes by Electroporation", Appl. Environ. Microbiol. vol. 61, pages 352-356, or by protoplast transformation.
40 *Streptomyces* are preferably transformed using protoplast transformation as described in Hopwood, et. al. (1985), "Genetic Manipulation of *Streptomyces*: A Laboratory Manual", The John Innes Foundation, Norwich, UK. *E. coli* strains are transformed using conventional transformation procedures as described in Maniatis, et al. (1982)

"Molecular Cloning: A Laboratory Manual" Cold Spring Harbor Laboratory, N.Y.

Plasmid pANT195 was introduced into *Streptomyces lividans* TK24 by protoplast transformation according to the procedures described in
5 Hopwood et. al (1985), "Genetic Manipulation of *Streptomyces*: A Laboratory Manual", The John Innes Foundation, Norwich, UK. 500 μ l of *Streptomyces lividans* TK24 protoplasts, were transformed with 10 μ l of plasmid DNA, about 0.5 μ g total, in 500 μ l of T buffer for two minutes. The reaction was stopped with 500 μ l of P buffer and the protoplasts
10 were pelleted twice in a microcentrifuge for 7 seconds each spin. The pellets were then resuspended in 100 μ l of P buffer and plated onto R2YE medium using a soft R2YE agar overlay with 50 μ g/ml of thiostrepton added 24 hours later. The transformed microorganisms were tested for their ability to carry out daunomycin C-14 oxidation.

15 *S. peucetius* ATCC 29050, *Streptomyces coelicolor* CH999 and *Streptomyces* sp. strain C5 were transformed with plasmids pANT195 and pANT849 by protoplast transformation.

Daunomycin C-14 Hydroxylase

The daunomycin C-14 hydroxylase encoded by *doxA* is a cytochrome
20 P450-type enzyme having a deduced Mr of 46,096. Daunomycin C-14 hydroxylase is a monooxygenase which inserts a single oxygen at carbon 14 on daunomycin. The daunomycin C-14 hydroxylase also appears to catalyze the two step oxidation at C-13 from methylene to hydroxyl to a keto functional group. Daunomycin C-14 hydroxylase also oxidizes
25 13-dihydrocarminomycin to carminomycin and 13-dihydrodaunomycin to doxorubicin.

The deduced amino acid sequence of daunomycin C-14 hydroxylase which is encoded by *doxA* of strain C5 is shown in Figure 3 and SEQ. ID. No. 5.

30 Preparation of Daunomycin C-14 Hydroxylase

Example A

The daunomycin C-14 hydroxylase was isolated and partially purified and subjected to spectrophotometric analysis. First, *S. lividans* TK24 strains containing plasmid pANT195 were grown in YEME
35 medium containing 10 μ g/ml thiostrepton for 48 hours at 30°C, harvested and washed by centrifugation and then broken in 100 mM, pH 7.5 sodium phosphate buffer using a French pressure cell at 15,000 lb/in². The cell debris and unbroken mycelia were pelleted by centrifugation at 10,000 \times g for 30 minutes at 4°C, after which the supernatant was
40 analyzed by visible spectrometry. The cytochromes within the supernatant derived from the cultures were reduced by a few grains of sodium dithionite. The supernatant samples were bubbled with carbon monoxide for 1 minute prior to analysis. Spectra were obtained using

-12-

a Beckman DU-64 single beam spectrophotometer and reduced-plus-CO minus reduced difference spectra were obtained by electronic subtraction.

Reduced-plus-CO minus reduced difference spectra of samples derived from cultures of *S. lividans* TK24 containing plasmid pANT195 revealed a peak at 450 nm, characteristic of cytochrome P450 enzymes. Such peak was not observed in samples derived from the control culture, which lack the *doxA* gene.

Proteins derived from both *S. lividans* TK24 containing plasmid pANT195 and the control culture *S. lividans* TK24 containing control plasmid pANT849, were visualized by sodium dodecylsulphate polyacrylamide gel electrophoresis. Samples derived from cultures containing plasmid pANT195 revealed a band with M_r of about 42,000, close to the predicted size of daunomycin C-14 hydroxylase. This band was not present in samples derived from the control cultures.

15 Example B

A 50 ml culture of *S. lividans* TK24 (pANT195) was grown for 48 hours in YEME medium plus 10 μ g/ml thiostrepton as in Example 1. This culture was split into 2 X 25ml aliquots, each of which was used to inoculate a 1000 ml flask containing 225 ml of YEME medium plus 10 μ g/ml thiostrepton, giving 2 fresh 250 ml cultures, which were grown as described in Example 1 for 48 hours. A 14-liter stirred tank fermentor containing 9.5 liters of YEME medium with 10 μ g/ml of thiostrepton was inoculated with both 250 ml cultures, a total inoculum size of 500 ml, and the 10 L culture was incubated for 6 days under the following conditions: temperature, 28°C; air flow, 1 volume air/volume culture/minute; agitation, 250 rpms. The culture was harvested by continuous centrifugation using a Heraeus 300 MD System, from Heraeus Sepatech, South Plainfield, NJ, at 15,000 rpm and a flow rate of 100 ml/min. The resultant pellet was frozen at -70°C until further use.

30 A small portion of the frozen pellet of *S. lividans* TK24 (pANT195) was thawed on ice in ice-cold 0.1 M sodium phosphate ($\text{Na}_2\text{HPO}_4:\text{NaH}_2\text{PO}_4$) buffer having a pH 7.5. The thawed suspension was passed twice through a 4°C French Pressure cell at 15,000 pounds per square inch to break the cells. The broken cell suspension was centrifuged at 10,000 x g for 30 minutes at 4°C and the supernatant from this centrifugation step was kept on ice to provide an isolated partially purified daunomycin C-14 hydroxylase.

-13-

P450 Determination

A 100 μ l aliquot of the daunomycin C-14 hydroxylase prepared according to Example B, was added to 900 μ l of 0.1M sodium phosphate buffer having a pH of 7.5, in a cuvette and approximately 1 mg of sodium dithionite was added to reduce the sample. This sample was used as the background for a spectrophotometric scan from 400 nm to 600 nm. Carbon monoxide was bubbled through this sample for one minute and the sample was scanned again from 400-600 nm. Electronic subtraction of the reduced plus carbon monoxide minus reduced sample revealed a sharp peak at 450 nm, indicative of the active cytochrome P450 enzyme. This assay was used before, during and after Examples 19 to 26 to ensure that the daunomycin C-14 hydroxylase was active and stable. In all cases, the cytochrome P450 activity appeared to be 100% of the original.

15 Method C

The Fusion Protein

Plasmid pANT199 was introduced by transformation into *E. coli* strain TOP10 from Invitrogen. Transformants were selected using ampicillin and grown in 3.0 ml cultures of SOB medium overnight at 37°C. The recipe for the SOB medium was provided by Invitrogen. Fifty μ L of this culture was used to inoculate 3.0 ml of fresh SOB medium. The new culture was grown at 37°C for 2 hours to an optical density of 0.6 and then induced with IPTG at 1.0 mM final concentration for 5 hours. The culture was then harvested by centrifugation in a microcentrifuge and the pellet was frozen overnight at -20°C. The next day the pellet was boiled in SDS-PAGE sample buffer described in Laemmli, U.K. (1970) "Cleavage of Structural Proteins during the assembly of the head of Bacteriophage T4" Nature volume 227, pages 680-685, and run on a 10% (w/v) SDS-PAGE gel. A protein with M_r of about 52,000 was observed that was insert-specific, the approximate size expected for the fusion protein based on amino acid sequence.

The fusion protein is then bound to a nickel-agarose column from Invitrogen, Inc., San Diego, Cal., and washed with 50 mM sodium phosphate buffer at pH 8.0 containing also 300 mM NaCl and 20 mM imidazole. The protein is then eluted using the same buffer but containing with 250 mM imidazole buffer at pH of 8, to provide a pure fusion protein with a modified N-terminus as shown in Figure 9 and SEQ. ID. NO. 8. The leader sequence is then cleaved from the fusion protein using enterokinase available from Biozyme Lab. Int'l Ltd. San Diego Cal., according to the manufacturer's directions, to provide pure N-terminal-modified daunomycin C14 hydroxylase as shown in SEQ. ID. NO. 9.

-14-

Methods of Converting Daunomycin to Doxorubicin

A host microorganism transformed with a plasmid containing the *doxA* gene is grown preferably in liquid culture, and daunomycin is added to the culture broth. Preferably, the daunomycin is added at a concentration of from about 2 mg/L to 22.2 mg/L, more preferably about 2 to 10 mg/L. Preferably, the daunomycin concentration is below about 10 mg/L. Where the concentration is above about 10 mg/L, the daunomycin tends to kill the host microorganisms although doxorubicin is still produced. The culture of transformed host microorganism is then incubated with the daunomycin; the longer the incubation the greater the amount of daunomycin is converted to doxorubicin. Preferably, the culture is incubated at least 6 hours, more preferably, at least 24 hours with the daunomycin.

A 48 hour culture has sufficient biomass to convert 2 mg/L daunomycin to doxorubicin within 24 hours.

Next, the doxorubicin is extracted preferably from both the transformed microorganisms and the culture fluid, using conventional techniques. A suitable technique involves extracting the transformed microorganisms and the culture fluid, preadjusted to a pH of about 8.5, with a mixture of chloroform and methanol and separating and drying the organic phase to provide a culture extract. The culture extract is resuspended in methanol and the components of the culture extract are separated, preferably by chromatography, to provide doxorubicin.

Media Composition

GPS production medium contains: glucose, 22.5 g/L; Proflo from Traders, Memphis, TN, 10 g/L; NaCl, 3 g/L; CaCO₃, 3 g/L, and 10 ml/L trace salts according to Dekleva, M.L. et. al. (1985), "Nutrient Effects on Anthracycline Production by *Streptomyces peucetius* in a Defined Medium", Canad. J. Microbiol. vol. 31, pages 287-294.

APM seed medium contains the following: yeast extract, 8 g/L; malt extract, 20 g/L; NaCl, 2 g/L 3-(N-morpholino)propanesulfonic acid buffer, 15 g/L; antifoam B from Sigma Chemical Co., St. Louis, MO, 4 ml/L; 10% weight to volume MgSO₄, 1 ml/L; 1% weight to volume FeSO₄, 1 ml/L; 10% weight to volume ZnSO₄, 0.1 ml/L; 50% weight to volume glucose, 120 ml/L, added after autoclaving; tap water to 1.0 L as described in Guilfoile and Hutchinson, (1991), "A Bacterial Analog of the *mdr* Gene of Mammalian Tumor Cells is present in *Streptomyces peucetius*, the Producer of Daunorubicin and Doxorubicin", Proc. Nat'l. Acad. Sci. USA volume 88, pages 8553-8557.

The YEME medium contained 3 g/L yeast extract available from U.S. Biochemical Corp. Cleveland, Ohio; 5 g/L bacto-peptone from Difco Detroit, MI; 3 g/L Difco malt extract; 10 g/L glucose; 200 g/L sucrose; and 2 ml/L of an autoclave-sterilized solution of 2.5 M MgCl₂·6H₂O. The pH was adjusted to 7.2 and the solution was autoclaved

-15-

at 121°C for 20 minutes at 15 psi, to provide the YEME medium. The nitrate-defined-plus-yeast extract medium, also referred to herein as "NDYE medium", contains the following: yeast extract, 5.0 g/L; N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid) buffer 4.8 g/L; 0.06 g/L anhydrous MgSO₄; 0.24 g/L K₂HPO₄·3H₂O; 4.28 g/L NaNO₃; 1.0 ml/L 20X trace elements; 45% (w/v) glucose solution, 50 ml/L; pH 7.3. The 20X trace elements solution contains the following elements in double distilled water: ZnCl₂, 800 mg/L; FeCl₃·6H₂O, 4000 mg/L; CuCl₂·2H₂O, 40 mg/L; MnCl₂·4H₂O, 40 mg/L; Na₂B₄O₇·10H₂O, 40 mg/L; (NH₄)₆Mo₇O₂₄·4H₂O, 200 mg/L; NiCl₂, 100 mg/L as described by Dekleva et al. (1985) Can. J. Microbiol volume 31, pages 287-294. A few drops of Mazu DF60-P antifoam, obtained from Mazer Chemical Co., Gurney, IL, were added to control foaming in cultures containing NDYE medium. Other antifoam agents, such as Sigma Antifoam B from Sigma Chemical Company, are also suitable.

EXAMPLES

The following examples are illustrative and not intended to be limiting.

Methods of Producing Doxorubicin Employing Host Microorganisms

20 Example 1

Cultures of *S. lividans* TK24 containing plasmid pANT195 and control cultures of *S. lividans* TK24 containing plasmid pANT849 were grown on R2YE agar medium containing 10 µg/ml of thiostrepton in standard 100 mm x 15 mm plastic petri dishes for 5 days at 30°C, at which time the entire cultures had sporulated. The spores from one entire petri plate were used to inoculate 50 ml of modified YEME medium in 250 ml erlenmeyer flasks. Then 10 µg/ml thiostrepton in DMSO were added to the medium. The thiostrepton was added as selective pressure to maintain the plasmids. The cultures were grown for 48 hours at 30°C with rotary shaking at 250 rpm, one inch throw on the shaker. After 48 hours, a 1.5 ml sample was removed for plasmid analysis to ensure the presence and size of the insert DNA containing the *doxA* gene using a "mini-prep" procedure according to Carter, M.J., and I.D. Milton, (1993), "An Inexpensive and Simple Method for DNA Purification on Silica Particles," Nucleic Acids Res., Vol. 21, page 1044. Restriction endonuclease digestion with BspEI and agarose gel electrophoresis generated 1.7 kbp, 2.2 kbp, and 3.14 kbp fragments which indicated an intact plasmid and insert. Then a 10 µL solution containing 100 µg of filter sterilized daun mycin-HCl in distilled water was added to the cultures for a final concentration of 2.0 µg of daunomycin per ml of culture broth. Incubation was continued for 72 hours.

-16-

After a total of 120 hours, that is, 48 hours growth and another 72 hours of continued growth in presence of the daunomycin, the pH of the culture was 7.6. The remaining culture broth, a volume of 48.5 ml, was brought to pH 8.5 with the dropwise addition of 5 N NaOH. Then the whole culture broth, including culture fluid and cells, was extracted once with a 2X volume of chloroform:methanol at a ratio of 9:1. The organic phase was separated from the aqueous phase by centrifugation at about 10,000 x g for 10 minutes and then the organic phase was removed by pipet. The organic phase was air dried in a chemical fume hood, then resuspended in 1 ml of 100% reagent grade methanol and spotted onto aluminum-backed, 0.25 mm silica gel thin layer chromatography plates from Whatman, Clifton, NJ. The components derived from the organic phase culture extract were separated using a solvent system of chloroform:methanol:acetic acid:water (80:20:16:6). The anthracyclines in the culture extracts were visualized on the plates by their normal pigmentation and by their fluorescence under ultraviolet irradiation at 365 nm. Table 1 shows the results.

The culture extracts of Example 1 and known standards also were separated and analyzed by high performance liquid chromatography using a C₁₈ µBondapak reverse phase column from Waters Corp. Milford, Ma. The solutions of standards and culture extracts of Example 1 were filtered through 0.2 µm Nylon Acrodisc[®] 13 filters from Gelman Sciences, Ann Arbor, MI and separated by HPLC using a mobile phase of methanol:water (65:35) brought to a pH of 2.5 with 85% phosphoric acid using a Waters 600E Multisolvant Delivery Pump and Controller and U6K 0-2.0 ml manual injector and detected on-line at 254 nm using a Waters 486 Tunable Absorbance Detector. The data were analyzed on-line and post-run using "Baseline 815" software and a 386 SX PC-compatible computer. The products extracted from the cultures were compared to standards run in parallel and by co-chromatography. These results are shown in Table 2.

TABLE 1
THIN LAYER CHROMATOGRAPHY OF DOXORUBICIN PRODUCED
ACCORDING TO EXAMPLE 1 AS COMPARED TO KNOWN STANDARDS

	Sample	R _f of Sample
35	Daunomycin Standard	0.56
	Doxorubicin Standard	0.36
	13-Dihydrodaunomycin Standard	0.39
	Doxorubicin from Cultures Containing Plasmid <i>doxA</i> (pANT195)	0.36
40	13-Dihydrodaunomycin from Control Culture	0.39

A few grains of each standard was reconstituted in 1 ml of methanol.

As indicated by the results in Table 1, the cultures transformed with plasmid pANT195 containing *doxA* which were incubated with

daunomycin, produced doxorubicin. In contrast, the control cultures produced 13-dihydrodaunomycin. Co-chromatography confirmed these results.

TABLE 2

HIGH PERFORMANCE LIQUID CHROMATOGRAPHY OF
DOXORUBICIN PRODUCED ACCORDING TO EXAMPLE 1
AS COMPARED TO KNOWN STANDARDS

Sample	Retention time of sample ² (minutes)
Daunomycin ¹ Standard	15.3
Doxorubicin ¹ Standard	8.4
13-Dihydrodaunomycin Standard	10.6
Doxorubicin from <i>doxA</i> Transformants	4.1; 7.2 (minor) 8.4 (major)
Control	4.1; 10.6; 21.2

¹ Standards were reconstituted at 1.0 mg/ml in methanol.

² A methanol peak at 2.9 minutes was found in all samples.

The *doxA* transformants converted greater than 90% of the daunomycin to doxorubicin in 72 hours as evidenced by both TLC and HPLC analyses. The control cultures, which lack the *doxA* gene, converted daunomycin to 13-dihydrodaunomycin, but not to doxorubicin.

The culture extracts of Example 1 and the standards were hydrolyzed to their respective aglycones to verify the chemical structures. The acid hydrolysis product of the doxorubicin is adriamycinone, and the acid hydrolysis product of the 13-dihydrodaunomycin is 13-dihydrodaunomycinone. The acid hydrolysis product of the doxorubicin produced by Example 1 was adriamycinone. The acid hydrolysis product of the 13-dihydrodaunomycin produced by control cultures of Example 1 was 13-dihydrodaunomycinone.

Example 2

The procedure of Example 1 was repeated, with the following exceptions. The cultures were grown at 28°C rather than 30°C, for 48 hours. After 48 hours of growth, 500 µg of daunomycin-HCl were added to the cultures for a final concentration of 10.0 µg/ml of daunomycin followed by further incubation for 36 hours, instead of 72 hours.

The culture broth then was extracted and the entire sample volume was spotted in a line onto a 250 µm layer thickness 20 cm x 20 cm glass-backed TLC plate containing a fluorescent indicator (254 nm) from Aldrich, Milwaukee, WI. The doxorubicin was separated from contaminants by chromatography for 2 hours using a mobile phase of chloroform:methanol:acetic acid:water (80:20:16:6), after which the silica gel containing the band having R_f , 0.3 - 0.4 was scraped from the

-18-

plate. The silica gel was extracted three times with about 1.5 ml of methanol each time. The methanol extracts were combined, filtered through a 0.2 μ m Nylon Acrodisc[®] 13 filter, and air-dried. The dried product was resuspended in 500 μ L of chloroform:methanol in a ratio of 9:1, back-extracted with an equal volume of water that had been previously made alkaline to pH 10.0 using Na₂CO₃, and the organic phase from this extraction procedure was removed and dried. The dried sample was resuspended in 500 μ L of methanol, from which 50 μ L were removed for HPLC and TLC analysis. HPLC analysis of this sample confirmed that doxA transformant cultures converted virtually all of the 500 μ g of daunomycin to doxorubicin.

The remainder was redried and subjected for mass spectrometry analysis. MS spectra were recorded on a SCIEX API III+ triple quadrupole mass spectrometer fitted with an atmospheric pressure chemical ionization source operating in a positive ion mode. MS spectra were acquired by scanning the first quadrupole (Q1), the results are shown in Table 3.

TABLE 3

MS ANALYSIS ON THE DOXORUBICIN PRODUCED ACCORDING TO EXAMPLE 2

Sample	Calculated	Average
	MW	M+1
Daunomycin standard	527.51	528.00
Doxorubicin standard	543.54	544.05
13-Dihydrodaunomycin standard	529.50	530.01
Doxorubicin from Transformants	----	544.00
Containing doxA		

The results of the MS analysis, shown in Table 3, indicate that the doxorubicin from doxA transformed cultures has an M+1 of 544.00, essentially the same value as obtained with the doxorubicin standard. The M+1 value of the doxorubicin produced by the doxA transformed culture was not similar to the M+1 values obtained with either daunomycin standard or the 13-dihydrodaunomycin standard.

Example 3

A 50 ml culture of *Streptomyces lividans* TK24 (pANT195) culture was prepared as in Example 1. This culture was grown for 48 hours at 28°C and then 25 ml were removed and used to inoculate 200 ml of YEME medium containing 10 µg/L of thiostrepton in a 1.0 L flask, having total, 225 ml of culture volume. After incubation for 48 hours at 28°C, 5.0 mg of daunomycin-HCl in 1000 µl of distilled water were added to the culture to give a final concentration of 22.2 µg/ml. A control culture of *S. lividans* TK24 containing plasmid pANT849 which lacks the *doxA* gene, was incubated in the presence of 100 µg of daunomycin-HCl. The cultures were incubated for 48 hours and then extracted as described in Example 1 and analyzed by HPLC and TLC. The doxorubicin which migrated in a broad band having an R_f of 0.3 - 0.4 was separated from contaminants by chromatography and prepared for MS analysis as in Example 2. The results are shown in Table 4.

The *doxA* transformed culture converted essentially all of the 5 mg of daunomycin to doxorubicin. Notably, the *doxA* transformed culture was virtually dead at the end of 48 hours, whereas the cultures of Examples 1 and 2 which received 2 µg/ml of daunomycin were fully viable. Nevertheless, even though the culture was eventually killed by the daunomycin, the culture converted essentially all of the daunomycin to doxorubicin.

HPLC analysis showed that the *doxA* transformed cultures converted greater than 95% of the daunomycin to doxorubicin. The control culture converted essentially 100% of the daunomycin to 13-dihydrodaunomycin, and did not produce doxorubicin.

TABLE 4

MASS SPECTROPHOTOMETRY ANALYSIS ON THE
DOXORUBICIN PRODUCED ACCORDING TO EXAMPLE 3

Sample	Calculated MW	Average M+1
Doxorubicin Standard	543.54	543.65
13-Dihydrodaunomycin Standard	529.50	529.85
Doxorubicin from Transformants Containing <i>doxA</i>	----	543.90
13-Dihydrodaunomycin from Control culture	----	529.15

The results of the MS analysis, shown in Table 4, indicate that the doxorubicin produced by cultures containing plasmid pANT195, has an M+1 of 543.90, essentially the same as obtained with doxorubicin standard. An MS-MS analysis was run on the parent 543.9 peak and is shown in Table 5.

TABLE 5
MASS SPECTROPHOTOMETRY ANALYSIS OF DOXORUBICIN
PRODUCED ACCORDING TO EXAMPLE 3

Sample	M+1	Major fragmentation
5 Doxorubicin Standard	543.65	489.90 ⁿ , 396.95, 378.70, 360.45 345.95 ⁿ 320.85 299.45 130.15
10 Doxorubicin produced according to Example 3	543.90	396.80, 378.90, 361.00, 130.20
13-Dihydrodaunomycin Standard	529.85	382.85, 365.30, 346.35, 320.85, 129.35, 113.10
15 Control culture extract	529.15	497.20 ⁿ , 482.00 ⁿ , 382.80, 364.60, 320.75, 305.75, 129.95

1 The product sample was significantly less concentrated than the
20 standard sample, leading to recovery of only the most abundant
fragmentation species. " - minor fragmentation species.

The MS-MS analysis on 543.90 peak from the doxorubicin produced
in Example 3 shows daughter peaks which are essentially identical to
the daughter peaks from the doxorubicin standard. Thus, the
doxorubicin produced by the transformants containing *doxA* has the M+1
25 and MS/MS fragmentation patterns of standard doxorubicin. Similarly,
the culture extract from the control culture had an M+1 and
fragmentation pattern similar to that of standard 13-dihydrodaunomycin.

Example 4

30 Fifty ml cultures of *Streptomyces lividans* TK24 (pANT195) and *S.*
lividans TK24 (pANT849) as a control were inoculated and prepared as in
Example 1 except that they were grown at 28°C for 48 hours. At that
time, 100 µg of daunomycin in 10 µL of distilled water was added to the
cultures for a final concentration of 2 µg/ml. The cultures were
further incubated for 24 hours. The cultures were then extracted as
35 described in Example 1 and subjected to HPLC analysis.

The transformed culture containing the *doxA* gene converted
greater than 95% of the daunomycin to doxorubicin within 24 hours.
The control culture converted approximately 100% of the daunomycin to
13-dihydrodaunomycin.

40 Example 5

The procedure of Example 4 was repeated except that 100 µg of
13-dihydrodaunomycin, rather than daunomycin, were added to the
cultures and the cultures were further incubated for 48 hours rather
than 24 hours.

-21-

In 48 hours, the culture containing plasmid *doxA* (pANT195) converted 100% of the 13-dihydrodaunomycin to doxorubicin. The control culture did not convert the 13-dihydrodaunomycin.

Example 6

5 Fifty ml cultures of *Streptomyces lividans* TK24 (pANT196) and *S. lividans* TK24 (pIJ702) as a control, were treated as in Example 4 except that the cultures were incubated with daunomycin for 72 hours.

The culture containing plasmid pANT196, which contains the *doxA* gene expressed from the *melC1* promoter, converted 20% of the daunomycin to doxorubicin. Thus the *melC1* promoter is less preferred than the *snpA* promoter. The control cultures converted 100% of the daunomycin to 13-dihydrodaunomycin.

Example 7

50 ml cultures of *Streptomyces lividans* TK24 (pANT192), having wild type and *snpA* promoters, *Streptomyces lividans* TK24 (pANT193), having wild type and *snpA* promoters, *Streptomyces lividans* TK24 (pANT194) which lacks the *snpA*-promoter and *snpR* activator gene, and *S. lividans* TK24 (pANT849) as a control, were prepared and analyzed prepared as in Example 4, except that the cultures were incubated in the presence of the daunomycin for 48 hours. The results are presented in Table 6.

TABLE 6

COMPARISON OF DIFFERENT PLASMIDS ON
PERCENT CONVERSION OF DAUNOMYCIN TO DOXORUBICIN

25 <u>Plasmid</u>	<u>Products</u>
pANT192	75% 13-dihydrodaunomycin/25% doxorubicin
pANT193	80% doxorubicin/20% 13-dihydrodaunomycin
30 pANT194	90% 13-dihydrodaunomycin/10% doxorubicin
pANT195	100% doxorubicin
pANT849	100% doxorubicin
(control)	100% 13-DHD

35 As shown in Table 6, the culture containing pANT192, which contains *doxA*, converted 25% of daunomycin to doxorubicin and 75% of daunomycin to the 13-dihydrodaunomycin. The culture containing plasmid pANT193, which contains *doxA*, converted 80% of daunomycin to doxorubicin and 20% of daunomycin to the 13-dihydrodaunomycin. The culture containing plasmid pANT194, which contains *doxA* but lacks the *snpA*-promoter and *snpR* activator, converted 10% of daunomycin to doxorubicin and 90% of daunomycin to the 13-dihydrodaunomycin. The

-22-

control culture converted the daunomycin to 13-dihydrodaunomycin, but not doxorubicin.

Example 8

Fifty ml cultures of *Streptomyces lividans* TK24(pANT195) and *S. lividans* TK24(pANT849) as a control, were prepared as in Example 4, except that the pH of the YEME medium was adjusted before inoculation using NaOH or HCl to provide an initial culture pH as shown in Table 7. The cultures were further incubated for 48 hours, rather than 24 hours. The results are shown in Table 7.

10

TABLE 7

EFFECT OF CULTURE PH ON DOXORUBICIN PRODUCTION

	Initial pH	Final pH	Percent Daunomycin bioconverted to Doxorubicin
15	6.0	7.0	50%
	6.5	7.6	90%
	7.0	7.9	100%
	7.5	7.0	100%
	8.0	---	No growth or bioconversion

20 The % conversion is approximate.

The culture containing plasmid pANT195 which contains *doxA*, and which was initially at pH of 7.0 or 7.5, converted 100% of the daunomycin to doxorubicin. The cultures containing plasmid *doxA* which were initially at pH 6.0 converted 50% of the daunomycin to doxorubicin. The cultures containing plasmid *doxA* which were initially at pH 6.5 converted 90% of the daunomycin to doxorubicin. Accordingly it is preferred that the transformed host cultures be grown at an initial pH of higher than 6.5.

Example 9

30 The procedure of Example 4 was repeated except that cultures were grown at either 22°C, 28°C, or 37°C and incubated with the daunomycin for 48 hours at such temperatures.

All three of the cultures containing plasmid *doxA* converted 100% of the daunomycin to doxorubicin.

35 Example 10

A 50 ml culture of *Streptomyces lividans* TK24(pANT195), inoculated and prepared as in Example 1, was grown at 28°C for 48 hours. At that time, the cultures were harvested by centrifugation at 10,000 × g in a high speed centrifuge, washed once with 100 mM 3-(N-morpholino)propanesulfonic acid buffer at pH 7.2. The cells from the
40 cultures were reconstituted in 5.0 ml of the 100 mM 3-(N-

-23-

morpholino)propanesulfonic acid buffer to give a final volume of 6.0 ml which included the volume of the packed cell mass, resulting in an approximately 8-fold concentration of the recombinant mycelia in buffer. Then 100 µg of daunomycin were added in 10 µL of distilled water for a final concentration of 16.7 µg/ml of daunomycin. A concentration of 16.7 µg/ml of daunomycin is toxic to the host. The culture was further incubated for 7.0 hours, after which it was extracted as described in Example 1 and subjected to HPLC analysis.

In 7 hours, the concentrated cultures containing the plasmid with a *doxA* insert converted about 25% of the daunomycin to doxorubicin.

Example 11

Five ml each of APM seed medium containing 10 µg/ml of thiostrepton were inoculated by loop from R2YE agar plates, containing 50 µg/ml thiostrepton of *S. peucetius* 29050 (pANT195) and *S. peucetius* 29050 (pANT849). Each culture was grown at 28°C for 48 hours. Fifty ml each of GPS "production" medium containing 10 µg/ml of thiostrepton were inoculated with 2.5 mls of seed culture grown in APM seed medium. The cultures were grown at 28°C for 48 hours as in Example 1. Then 100 µg of daunomycin in 10 µL of distilled water for a final concentration of 2 µg/ml of daunomycin was added to the cultures. The cultures were further incubated for 48 hours, then extracted as described in Example 1.

After 48 hours, the culture containing plasmid pANT195, which contains the *doxA* gene, converted about one-half of the daunomycin to doxorubicin. The control cultures did not convert daunomycin to doxorubicin.

Example 12

Fifty ml cultures of *Streptomyces coelicolor* CH999 (pANT195), and *Streptomyces coelicolor* CH999 (pANT849) as a control, were used in the procedure of Example 4 except that the cultures were incubated with the daunomycin for 48 hours.

Again the cultures containing pANT195 converted 100% of the daunomycin to doxorubicin, while the control converted 80% of the daunomycin to 13-dihydrodaunomycin. In the control cultures 20% of the daunomycin was not converted.

Example 13

The procedure of Example 12 was repeated, except that after the daunomycin addition, the cultures were only incubated for 1, 2, or 4 hours.

After 1 hour, 0.7% of the daunomycin was converted to doxorubicin by the cultures which contained pANT195. After 2 hours, 1.0% of the daunomycin was converted and by 4 hours 15% of the daunomycin was converted to doxorubicin. The control cultures which

-24-

lacked the plasmid containing *doxA* did not convert any of the daunomycin.

Example 14

50 ml cultures of *Streptomyces* sp. strain C5 (pANT195) and
5 *Streptomyces* sp. strain C5 (pANT849) as a control, were grown at 28°C
for 72 hours in NDYE medium as described in Example 1. After 72 hours,
100 µg of unlabelled daunomycin and 5 µCi of ³H-daunomycin, having a
specific radioactivity of 5.0 Ci/mmol, were added to each culture and
they were incubated for another 48 hours. The cultures were extracted
10 as described in Example 1 and analyzed by TLC and autoradiography.

The cultures of containing plasmid pANT195 converted
approximately 5% of the radiolabelled daunomycin to doxorubicin. No
other products other than the substrate daunomycin and doxorubicin were
observed in these cultures. The control cultures converted
15 approximately 90% of the radiolabelled daunomycin to baumycin A1 and
baumycin A2.

Example 15

50 ml cultures of *Streptomyces* sp. strain C5, which does not
appear to synthesize doxorubicin, and the following mutants of
20 *Streptomyces* sp. strain C5: SC5-dauA74, SC5-dauCE147, SC5-dauE24, and
SC5-dauH54, were grown for 48 hours in NDYE medium. These mutants do
not synthesize daunomycin. At 48 hours, 100 µg of daunomycin was
added for a final concentration of 2 µg/ml, and the cultures were
incubated for 36 hours. The products were extracted as in Example 1
25 and subjected to HPLC analysis.

All *Streptomyces* sp. strain C5 cultures, each of which lacked a
plasmid containing *doxA*, converted about 10% of the daunomycin to
doxorubicin. Baumycins A1 and A2 were also detected.

Example 16

30 The procedure of Example 4 was repeated except that the cultures
were then incubated for 48 hours with 100 µg, for a final concentration
of 2 µg/ml of one of the following: carminomycin, idarubicin,
daunomycinone, or carminomycinone.

Incubation of *S. lividans* TK24 (pANT195) cultures with
35 carminomycin or idarubicin, resulted in greater than 85% recovery of
carminomycin and idarubicin. The cultures which contained plasmid
pANT195 converted 100% of the daunomycinone to 13-hydroxydaunomycinone
and 100% of the carminomycinone to 13-hydroxycarminomycinone. The
control cultures, which lack the *doxA* gene, converted 100% of the
40 carminomycin, idarubicin, daunomycinone, and carminomycinone to their
13-dihydro derivatives.

Example 17

The procedure of Example 4 was repeated except that the cultures
received 100 µg of 13-dihydrocarminomycin rather than daunomycin. The

-25-

cultures were then incubated for 36 hours and analyzed by TLC and HPLC.

The cultures containing plasmid pANT195 converted 100% of the 13-dihydrocarminomycin to carminomycin. No other products were observed. In the control culture, none of the 13-dihydrocarminomycin was converted. The *doxA* gene confers the ability to oxidize the C-13 hydroxyl function of the 13-dihydro-carminomycin to a keto function.

Novel Synthesis of 13-deoxycarminomycin and 13-deoxydaunomycin

Example 18

10 ϵ -Rhodomycin D compound was converted to both 13-deoxycarminomycin and 13-deoxydaunomycin, by host microorganisms containing plasmid which contains the *Streptomyces* sp. strain C5 *dauP* gene which encodes ϵ -rhodomycin D esterase and the *Streptomyces* sp. strain C5 *dauK* gene which encodes carminomycin 4-O-methyltransferase.

15 Plasmid pANT144 as shown in Figure 12, and described in Dickens, M.L., et. al., (1995) "Analysis of Clustered Genes Encoding both Early and Late Steps in Daunomycin Biosynthesis by *Streptomyces* sp. strain C5" *J. Bacteriol.* volume 177, pages 536-543, was introduced into *S. lividans* TK24 by protoplast transformation. *S. lividans* TK24 (pANT144) was grown

20 for 48 hours in 50 ml of YEME medium containing 10 μ g/ml of thiostrepton and then used to inoculate 450 ml of the YEME medium which contained 10 μ g/ml of thiostrepton for a total culture volume of 500 ml, in a two liter flask. The resultant 500 ml culture was incubated for 48 hours at 28°C as in Example 1. Next, 5.0 mg of ϵ -rhodomycin D,

25 the glycone of ϵ -rhodomycinone, from the National Cancer Institute, Drug Synthesis and Chemistry Branch, Bethesda, MD designated compound #263854-H, were added to the culture for a final concentration of 10 μ g/ml and the cultures were incubated for an additional 48 hours. The culture then was adjusted to pH 8.5, and then extracted twice, each

30 with 1 volume of chloroform:methanol (9:1) as in Example 3. The organic extract was reduced to dryness, reconstituted in 500 μ L of chloroform:methanol (9:1), filtered, back extracted, dried, and reconstituted in 2.0 ml of methanol. The extract was separated and extracted as in Example 2. R_f values for 13-deoxycarminomycin and 13-deoxydaunomycin were approximately 0.60 and 0.64, respectively. The 13-deoxycarminomycin and 13-deoxydaunomycin were reduced to dryness and each was brought up again in 50 μ l of methanol.

Methods of Producing Anthracyclines Employing Daunomycin C-14 Hydroxylase

40 Example 19

904 μ l of the daunomycin C-14 hydroxylase produced according to the method of Example B, containing approximately 1 mg of total

protein, was incubated in a 16 mm well of a 24 well culture plate at 30°C with shaking for 2 hours with 25 µg in 5 µl of either daunomycin, 13-dihydrodaunomycin, or 13-dihydrocarminomycin. The final volume of each well was 1.0 ml; 0.1M sodium phosphate buffer at pH 7.5, was added to bring the total volume to 1 ml, as needed. After 2 hours of incubation, the pH of each reaction mixture was increased to pH 8.5 using 1 M NaOH and each was extracted twice each with 500 µl of chloroform: methanol (9:1). The organic layers were combined, reduced to dryness, reconstituted in 10 µl of methanol, separated and analyzed by TLC as in Example 1.

The daunomycin C-14 hydroxylase converted 50% of the 13-dihydrocarminomycin to carminomycin and 50% of the 13-dihydrodaunomycin to daunomycin.

Example 20

The procedure of example 19 was repeated except that 10 µl of NADH, 1 mM final concentration; 10 µl of NADPH, 1 mM final concentration, were added.

The daunomycin C-14 hydroxylase converted 100% of the 13-dihydrocarminomycin to carminomycin and 100% of the 13-dihydrodaunomycin to daunomycin.

Example 21

The procedure of Example 20 was repeated except that the following reagents, available from Sigma Chemical Co., were added: 20 µl of glucose-6-phosphate; 10 mM final concentration; 10 µl of NADP⁺, 1 mM final concentration; 1.0 µl of glucose-6-phosphate dehydrogenase; 0.84 units, final activity; 20 µl of spinach ferredoxin, 44 µg final concentration; and 10 µl of spinach ferredoxin-NADP⁺ reductase, 0.05 units final activity. The glucose-6-phosphate, glucose-6-phosphate dehydrogenase and NADP⁺ constitute a "NADPH-regenerating system". After 2 hours of incubation, the extracts were extracted and analyzed by TLC and HPLC.

The daunomycin C-14 hydroxylase converted 100% of the 13-dihydrocarminomycin to carminomycin and 100% of the 13-dihydrodaunomycin to daunomycin, as shown by TLC. HPLC revealed the conversion of about 5% of the daunomycin to doxorubicin.

Example 22

The procedure of Example 21 was repeated, except 10 µl of flavin adenine mononucleotide from Sigma, 10 µg final concentration; and 10 µg in 10 µl, flavin adenine dinucleotide, were also added.

100% of the 13-dihydrocarminomycin was converted to carminomycin and 100% of the 13-dihydrodaunomycin was converted to daunomycin by the daunomycin C-14 hydroxylase. Doxorubicin was not detected using TLC.

Example 23

-27-

The procedure of Example 20 was repeated except that different anthracyclines were used, incubation was for 1 hour and 10 μ l of NADP⁺, 1 mM final concentration was added. 25 μ g in 5 μ l the following anthracyclines were used: either ϵ -rhodomycin D, 13-deoxydaunomycin from example 18, or 13-deoxycarminomycin, from example 18.

About 20% of the 13-deoxycarminomycin was converted to 13-dihydrocarminomycin and about 80% was converted to carminomycin; and 13-deoxydaunomycin was converted to about 20% 13-dihydrodaunomycin and about 80% daunomycin. The ϵ -rhodomycin D did not appear to be converted.

Thus, the daunomycin C-14 hydroxylase converted 13-deoxycarminomycin to 13-dihydrocarminomycin and carminomycin; 13-dihydrocarminomycin to carminomycin; 13-deoxydaunomycin to 13-dihydrodaunomycin and daunomycin; and 13-dihydrodaunomycin to daunomycin. Thus, daunomycin C-14 hydroxylase catalyzes the oxidation of the C-13 methylene to a C-13 hydroxyl function, and catalyzes the oxidation of the C-13 hydroxyl function to C-13 keto function. The daunomycin C-14 hydroxylase is useful for making 13-dihydrocarminomycin, carminomycin, 13-dihydrodaunomycin and daunomycin.

Example 24

The procedure of Example 22 was repeated except that the daunomycin C-14 hydroxylase was incubated with daunomycin for 18 hours rather than 2 hours and in 25 ml erylenmyer flasks shaken at 250 rpm on a rotary shaker.

Doxorubicin was not detected by HPLC or TLC.

Example 25

The procedure of Example 24 was repeated except that the reagent volumes were tripled. The reagent concentrations however were not increased.

Approximately 5% of the daunomycin was converted to doxorubicin as determined by HPLC.

Example 26

The procedure of Example 24 was repeated except that the reagent volumes were quintupled. The reagent concentrations however, were not increased.

Approximately 20 to 25% of the daunomycin was converted to doxorubicin, as determined by HPLC.

The present invention includes: the DNA sequences encoding a protein daunomycin C-14 hydroxylase, which adds a hydroxyl group to carbon 14 of daunomycin; the messenger RNA transcript of such DNA sequence; and an isolated protein which adds a hydroxyl group to carbon 14 of daunomycin.

-28-

For example, the DNA sequences include: DNA molecules which, but for the degeneracy of the genetic code would hybridize to DNA encoding the daunomycin C-14 hydroxylase, thus the degenerate DNA which encodes the daunomycin C-14 hydroxylase protein; DNA strands
5 complementary to: DNA sequences encoding the daunomycin C-14 hydroxylase protein including DNA in Figures 3, 6, 9 and 11; heterologous DNA having substantial sequence homology to the DNA encoding the daunomycin C-14 hydroxylase protein, including the DNA sequences in Figures 3, 6, 9 and 11 or portions thereof.

10 The daunomycin C-14 hydroxylase protein includes, for example, the daunomycin C-14 hydroxylase protein of strains other than *Streptomyces* sp. strain C5; proteins having 75% homology to the proteins in Figures 3, 6, 9 and 11, and proteins or portions thereof having substantially the same amino acid sequence as shown in Figures
15 3, 6, 9 and 11.

-29-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Strohl, William R.
Dickens, Michael L.
DeSanti, Charles L.
- (ii) TITLE OF INVENTION: METHOD OF PRODUCING DOXORUBICIN
- (iii) NUMBER OF SEQUENCES: 9
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: CALFEE, HALTER & GRISWOLD
 - (B) STREET: 800 Superior Avenue, Suite 1400
 - (C) CITY: Cleveland
 - (D) STATE: Ohio
 - (E) COUNTRY: USA
 - (F) ZIP: 44114-2688
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Golrick, Mary E.
 - (B) REGISTRATION NUMBER: 34829
 - (C) REFERENCE/DOCKET NUMBER: 22727/00131
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 216-622-8458
 - (B) TELEFAX: 216-241-0816

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GACATGCATG CCGAGGGGTG CCTC
24

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid

-30-

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GACGCAGCTC CGGAACGGGG
20

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 48 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCATGCGAAT TCAGATCTAG AGCTCAAGCT TTAACTAGT TAACGCGT
48

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3196 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 1498..2764

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1498..2764

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGTACCCGCG CATCGATGTC ATGGCCGCGA ACGCCGGCGG CATGTTCTGG TCGCGCACCA	60
CGACCCAGGA CGGGTTCGAG GCCACCCTCC AGGTCAATCA TCTCGCGGGC TTCCTGCTGG	120
CACGGCTGCT GCGGGAGCGG CTCGCGGGCG GCGGGTTGAT CCTCACCTCG TCCGACGCGT	180
ACACCCAGGG CCGGATCGAC CCGACGATC TCAACGGCGA CCGTCACCGC TACAGCGCGG	240
GCCAGGCGTA CGGCACGTCC AAACAGGCCA ACATCATGAC CGCCACGGAG GCCGCCCCGGC	300
GCTGGCCGGA CGTGCTGACG GTCAGCTACC ACCCGGGCGA GGTCCGCACC CGCATCGGGC	360
GGGGCACAGT CGCCTCGACC TACTTCCGGT TCAACCCCTT CCTGCGGTCC GCGGCCAAGG	420

GCGCCGACAC TCTCGTGTGG CTGGCGGCCG CGCCGGCCGA GGAGTTGACC ACGGGCGGCT	480
ACTACAGCGA CCGGCGGCTG TCCCCGGTGA GCGGCCCGAC CGCCGACGCC GGCCTCGCGG	540
CCAAGCTCTG GGAGGCCAGC GCGGCCGCCG TCGGCCACAC CGCGCGCTGA CCGCGGCGGG	600
CCTCCCCGCC CGCATGCCCC TCTCATCCGC GAGCGCAGAC GCTCGTGTGC CGATCCGTGC	660
AAAGGAACGA TTCGTGACCA GGTTCGCGCC CGGCGCCCCC GCATGGTTCG ACCTCGGGTC	720
GCCCCGATGC GCCGCCTCGG CCGACTTCTA CACCGGCCCTC TTCGCGTGGA CCGCGACCGT	780
GGTCAGCGAC CCGGGTGCCG GGGGATACAC TACTTTTCAGC TCCGACGGGA AGCCTGTCCG	840
CGCGGTCGCC CGCCATCAGA TCGACACGCC CTACCACCGT CCGTACGGGC CCGCAAGCA	900
CCAGCACGGC ATGCCGGCCA TCTGGACCGT GTACTTCGCC ACCAACGACG CCGACGCACT	960
GACCAAACGG GTCGAAGCGG CCGGTGGCGA CGTCATCATG CACCCGATGG ACGTCTCGG	1020
TCTCGGCCGG ATGGCGTCT TCGCCGACCC ATCGGGGGCC GCGTTCGCGG TGTGGCGCAA	1080
GGGCGTCATG GAGGGCGCGG AGGTGACGGG CGTGCCCGGC TCGGTCGGCT GGGTGGAAC	1140
GGTGACCGAC GACATCGGGA CCGCCCGTGG CTTCTACCGT GCGACCCTCG GCCTGGCTCC	1200
GGCCGACACC GGACGCAAGG GCGTCACCGA CCCGGTTTGG CACATCCATG ACACACCGGT	1260
CGCCGGCACC CGGGAACCTG GCACGACCGG CGCGGTACGG CCCCCTGGG CCGTGCTGTT	1320
CTCCGTGCAC GACTGCGACG CGACGGTCCG GCGGGCCGTC GAACTCGGCG GCTCCGTCGA	1380
GAACGAGCCC GTCGACACCC CCAGGGGGCG GCGGGCGGAC CTGCTCGACC CGCACGGGGC	1440
CGGCTTCTCG GTGGTCGAAC TGCGGGAGGC GTACCCCGCG GCGGCGGACG GTGCCTC	1497
ATG AGC GGC GAG GCG CCG CGG GTG GCC GTC GAC CCG TTC TCG TGT CCC Met Ser Gly Glu Ala Pro Arg Val Ala Val Asp Pro Phe Ser Cys Pro	1545
1 5 10 15	
ATG ATG ACC ATG CAG CGC AAA CCC GAG GTG CAC GAC GCA TTC CGA GAG Met Met Thr Met Gln Arg Lys Pro Glu Val His Asp Ala Phe Arg Glu	1593
20 25 30	
GCG GGC CCC GTC GTC GAG GTG AAC GCC CCC GCG GGC GGA CCC GCC TGG Ala Gly Pro Val Val Glu Val Asn Ala Pro Ala Gly Gly Pro Ala Trp	1641
35 40 45	
GTC ATC ACC GAT GAC GCC CTC GCC CGC GAG GTG CTG GCC GAT CCC CGG Val Ile Thr Asp Asp Ala Leu Ala Arg Glu Val Leu Ala Asp Pro Arg	1689
50 55 60	
TTC GTG AAG GGA CCC GAT CTC GCG CCC ACC GCC TGG CGG GGG GTG GAC Phe Val Lys Gly Pro Asp Leu Ala Pro Thr Ala Trp Arg Gly Val Asp	1737
65 70 75 80	
GAC GGT CTC GAC ATC CCC GTT CCG GAG CTG CGT CCG TTC ACG CTC ATC Asp Gly Leu Asp Ile Pro Val Pro Glu Leu Arg Pro Phe Thr Leu Ile	1785
85 90 95	
GCC GTG GAC GGT GAG GAC CAC CGG CGT CTG CGC CGC ATC CAC GCA CCG Ala Val Asp Gly Glu Asp His Arg Arg Leu Arg Arg Ile His Ala Pro	1833
100 105 110	
GCG TTC AAC CCG CGC CGG CTG GCC GAG CGG ACG GAT CGC ATC GCC GCC Ala Phe Asn Pro Arg Arg Leu Ala Glu Arg Thr Asp Arg Ile Ala Ala	1881

-32-

115	120	125	
ATC GCC GAC CGG CTG CTC ACC GAA CTC GCC GAC TCC TCC GAC CGG TCG Ile Ala Asp Arg Leu Leu Thr Glu Leu Ala Asp Ser Ser Asp Arg Ser 130 135 140			1929
GGC GAA CCG GCC GAG CTG ATC GGC GGC TTC GCG TAC CAC TTC CCG CTG Gly Glu Pro Ala Glu Leu Ile Gly Gly Phe Ala Tyr His Phe Pro Leu 145 150 155 160			1977
TTG GTC ATC TGC GAA CTG CTC GGC GTG CCG GTC ACC GAT CCG GCA ATG Leu Val Ile Cys Glu Leu Leu Gly Val Val Thr Asp Pro Ala Met 165 170 175			2025
GCC CGC GAG GCC GTC GGC GTG CTC AAG GCA CTC GGC CTC GGC GGC CCG Ala Arg Glu Ala Val Gly Val Leu Lys Ala Leu Gly Leu Gly Gly Pro 180 185 190			2073
CAG AGC GCC GGC GGT GAC GGC ACG GAC CCT GCC GGG GAC GTG CCG GAC Gln Ser Ala Gly Gly Asp Gly Thr Asp Pro Ala Gly Asp Val Pro Asp 195 200 205			2121
ACG TCG GCG CTG GAG AGC CTT CTC CTC GAA GCC GTG CAC GCG GCC CGG Thr Ser Ala Leu Glu Ser Leu Leu Leu Glu Ala Val His Ala Ala Arg 210 215 220			2169
CGG AAA GAC ACC CGG ACC ATG ACC CGC GTG CTC TAT GAA CCG GCA CAG Arg Lys Asp Thr Arg Thr Met Thr Arg Val Leu Tyr Glu Arg Ala Gln 225 230 235 240			2217
GCA GAG TTC GGC TCG GTC TCC GAC GAC CAG CTC GTC TAC ATG ATC ACC Ala Glu Phe Gly Ser Val Ser Asp Asp Gln Leu Val Tyr Met Ile Thr 245 250 255			2265
GGA CTC ATC TTC GCC GGC CAC GAC ACC ACC GGC TCG TTC CTG GGC TTC Gly Leu Ile Phe Ala Gly His Asp Thr Thr Gly Ser Phe Leu Gly Phe 260 265 270			2313
CTG CTT GCG GAG GTC CTG GCG GGC CGT CTC GCG GCG GAC GCC GAC GGG Leu Leu Ala Glu Val Leu Ala Gly Arg Leu Ala Ala Asp Ala Asp Gly 275 280 285			2361
GAC GCC ATC TCC CGG TTC GTG GAG GAG GCG CTG CCG CAC CAC CCG CCG Asp Ala Ile Ser Arg Phe Val Glu Glu Ala Leu Arg His His Pro Pro 290 295 300			2409
GTG CCC TAC TCG TTG TGG AGG TTC GCT GCC ACG GAG GTG GTC ATC CGC Val Pro Tyr Ser Leu Trp Arg Phe Ala Ala Thr Glu Val Val Ile Arg 305 310 315 320			2457
GGT GTC CCG CTG CCC CGC GGA GCG CCG GTA CTG GTG GAC ATC GAG GGC Gly Val Arg Leu Pro Arg Gly Ala Pro Val Leu Val Asp Ile Glu Gly 325 330 335			2505
ACC AAC ACC GAC GGC CGC CAT CAC GAC GCC CCG CAC GCT TTC CAC CCG Thr Asn Thr Asp Gly Arg His His Asp Ala Pro His Ala Phe His Pro 340 345 350			2553
GAC CGC CCT TCG AGG CGG CGG CTC ACC TTC GGC GAC GGG CCG CAC TAC Asp Arg Pro Ser Arg Arg Arg Leu Thr Phe Gly Asp Gly Pro His Tyr 355 360 365			2601
TGC ATC GGG GAG CAG CTC GCC CAG CTG GAA TCG CGC ACG ATG ATC GGC Cys Ile Gly Glu Gln Leu Ala Gln Leu Glu Ser Arg Thr Met Ile Gly 370 375 380			2649

-33-

GTA CTG CGC AGC AGG TTC CCC CAA GCC CGA CTG GCC GTG CCG TAC GAG	2697
Val Leu Arg Ser Arg Phe Pro Gln Ala Arg Leu Ala Val Pro Tyr Glu	
385 390 395 400	
GAG TTG CGG TGG TGC AGG AAG GGG GCC CAG ACA GCG CGG CTC ACT GAC	2745
Glu Leu Arg Trp Cys Arg Lys Gly Ala Gln Thr Ala Arg Leu Thr Asp	
405 410 415	
CTG CCC GTC TGG CTG CGT T GATGGGCCGA CCGCGACCCG GCACGGGACC	2794
Leu Pro Val Trp Leu Arg	
420	
GGCCACCGCC CATCGCGCGG TGGGCGGTCC CGTGCCGGTC GCCCGGTGCG GTCCTCTCCC	2854
GACGCTCGCT CCCCTGTGA CTTTCTCACA TCGAGACGTG ACGAAATAAT CCCAGCAAGT	2914
GCCATGCACA CTTTCATGGC GGACATTCACT TTGCGAGGAT GGAGTGAGCA CACGGGGCCG	2974
CCCGAGACAC CCTACGGCCG CCGGAAGTAT GCCACCTGTT GACGCGAATG GAACGCCACA	3034
GAGGGAGCAC CGGCAATGCA GATCAATATG TTGGGCCCCG TC GTTGACACA TCACAATGGC	3094
ACGTCGGTGA CCCCAGATAGC CAGAAAACCC CGGCAGGTAT TCTCACTGCT CGCTCTTCAG	3154
GCAGGAACCG TCGTTCCGGT CCCCAGCGTG ATGGAGGAGC TC	3196

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Ser	Gly	Glu	Ala	Pro	Arg	Val	Ala	Val	Asp	Pro	Phe	Ser	Cys	Pro
1				5					10					15	
Met	Met	Thr	Met	Gln	Arg	Lys	Pro	Glu	Val	His	Asp	Ala	Phe	Arg	Glu
			20					25					30		
Ala	Gly	Pro	Val	Val	Glu	Val	Asn	Ala	Pro	Ala	Gly	Gly	Pro	Ala	Trp
		35					40					45			
Val	Ile	Thr	Asp	Asp	Ala	Leu	Ala	Arg	Glu	Val	Leu	Ala	Asp	Pro	Arg
	50					55					60				
Phe	Val	Lys	Gly	Pro	Asp	Leu	Ala	Pro	Thr	Ala	Trp	Arg	Gly	Val	Asp
65					70					75				80	
Asp	Gly	Leu	Asp	Ile	Pro	Val	Pro	Glu	Leu	Arg	Pro	Phe	Thr	Leu	Ile
			85					90						95	
Ala	Val	Asp	Gly	Glu	Asp	His	Arg	Arg	Leu	Arg	Arg	Ile	His	Ala	Pro
		100						105					110		
Ala	Phe	Asn	Pro	Arg	Arg	Leu	Ala	Glu	Arg	Thr	Asp	Arg	Ile	Ala	Ala
		115					120					125			
Ile	Ala	Asp	Arg	Leu	Leu	Thr	Glu	Leu	Ala	Asp	Ser	Ser	Asp	Arg	Ser
	130					135					140				
Gly	Glu	Pro	Ala	Glu	Leu	Ile	Gly	Gly	Phe	Ala	Tyr	His	Phe	Pro	Leu

-34-

145					150					155					160
Leu	Val	Ile	Cys	Glu	Leu	Leu	Gly	Val	Pro	Val	Thr	Asp	Pro	Ala	Met
				165					170						175
Ala	Arg	Glu	Ala	Val	Gly	Val	Leu	Lys	Ala	Leu	Gly	Leu	Gly	Gly	Pro
			180					185					190		
Gln	Ser	Ala	Gly	Gly	Asp	Gly	Thr	Asp	Pro	Ala	Gly	Asp	Val	Pro	Asp
		195					200					205			
Thr	Ser	Ala	Leu	Glu	Ser	Leu	Leu	Leu	Glu	Ala	Val	His	Ala	Ala	Arg
		210				215					220				
Arg	Lys	Asp	Thr	Arg	Thr	Met	Thr	Arg	Val	Leu	Tyr	Glu	Arg	Ala	Gln
225					230					235					240
Ala	Glu	Phe	Gly	Ser	Val	Ser	Asp	Asp	Gln	Leu	Val	Tyr	Met	Ile	Thr
				245					250					255	
Gly	Leu	Ile	Phe	Ala	Gly	His	Asp	Thr	Thr	Gly	Ser	Phe	Leu	Gly	Phe
			260					265					270		
Leu	Leu	Ala	Glu	Val	Leu	Ala	Gly	Arg	Leu	Ala	Ala	Asp	Ala	Asp	Gly
		275					280					285			
Asp	Ala	Ile	Ser	Arg	Phe	Val	Glu	Glu	Ala	Leu	Arg	His	His	Pro	Pro
	290					295					300				
Val	Pro	Tyr	Ser	Leu	Trp	Arg	Phe	Ala	Ala	Thr	Glu	Val	Val	Ile	Arg
305					310					315					320
Gly	Val	Arg	Leu	Pro	Arg	Gly	Ala	Pro	Val	Leu	Val	Asp	Ile	Glu	Gly
				325					330					335	
Thr	Asn	Thr	Asp	Gly	Arg	His	His	Asp	Ala	Pro	His	Ala	Phe	His	Pro
			340					345					350		
Asp	Arg	Pro	Ser	Arg	Arg	Arg	Leu	Thr	Phe	Gly	Asp	Gly	Pro	His	Tyr
		355					360					365			
Cys	Ile	Gly	Glu	Gln	Leu	Ala	Gln	Leu	Glu	Ser	Arg	Thr	Met	Ile	Gly
	370					375					380				
Val	Leu	Arg	Ser	Arg	Phe	Pro	Gln	Ala	Arg	Leu	Ala	Val	Pro	Tyr	Glu
385					390					395					400
Glu	Leu	Arg	Trp	Cys	Arg	Lys	Gly	Ala	Gln	Thr	Ala	Arg	Leu	Thr	Asp
				405					410						415
Leu	Pro	Val	Trp	Leu	Arg										
				420											

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3013 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCAGGCGGTA CCGCCGACCC GCTGCATCCC CCGCACCGCC GTCCCCCCCC AGGGCATCTC	60
CCGTCGGGTT ACGGGAAGGG GGCCGGGGTA CCCGGTCGTC ACGGGAGGGC TGGGACGAGT	120
GCCCCGACC CACTGCGTTC CAGCCACTCC CGGTACGCCG GGGCCTGCCG GGCGACCTCC	180
CCGTAGGCCT CCGCGAGGTC GGGGTAGACG CCCTCCAGTT CCGTGTCCGT GCGGGCCGCC	240
AGCAGCAGCC GTACGCCGAG CCGGTCGCCG TGCAGCCGGC GGACGGCCGT CTCGGCGCGG	300
GAGGGCGAGG TCGGCTGGAC CACGGTGACG ACCTCGCCGG TGGCGACCAG GTACGCGGCG	360
GAGTGGTAGT CCCCCTGCAG GATGCGCGAG TCGAGTCCCT CGGCGCGCAG GACCCGGCGC	420
ACCGCGTTCC ACTCGCCGTC GACGGTGGGG TCGATCATCC AGCGGGTCGT GGGCCAGGTC	480
GGCGAGGCGT ACGACGTGGC TTCGGCGGCC GGGTGGTCGG CCGGCAGGCA GACGAACTGC	540
GGTTCCCGCT GGACCACTAC GCGGACCCGG AGCCCTTCGG GGACGCGCAG GCTGCCCTCG	600
ACCTCGTGCA CGAAGGCGAC GTCGAGGTGG CCGTCGGCCA CCATGCGCAG CAGGGCGTTG	660
GCGGAGACGT CCATGTGCAG GGTGGGTTCC TGCCAGTGCC GGAGCCGGCG CAGCCAGCCC	720
GCCAGGGCCC GGCTGGCCGT GGAGCCGACG CGCAGGCTGG CGTCCGCGAC GGCGGCGGCG	780
CGGGCCTCGC TGACGAGGGA GCACAATTCG GCCACCAGGG GGCGGGCAGC ACTGAGAACC	840
AGCCGGCCCA GCGGTGTGGG GCGGCAGCCG GTGCGGGCCC GGACGAACAG GGCACGGCCC	900
AGCTCGTGTT GGATGCGCCC CAGCTGCGTG CTCACGAGG GCTGTGTAC TCCCAGTTGG	960
CGTGCCGCGC GGTGCAGGCT GCCGGTGTGG GCGATGGCGC ACAGCGCCCT GAGGTGCCTG	1020
ACCTCAAGCT CCATGTCCTG GGAGGGTAAG GCGGAAGTTC AGCTTTCACC AGACATACAA	1080
AATGGCGACC GATCAGGACC ATCGGGCCTT CACGGCGCGA GGCCTCGGCC CGGATCGGCA	1140
GGGGCCCCGG CCGGGGCCGC CGGGCAGGGC GCGCAGGTG GGGACGGAGG GGGATAGGGC	1200
GGCCCTATCG GCGGTTGCCA TCATCACAAC GGCCGTACGG GCACGGACAC TCACGATGTC	1260
TGACTCATCC CCCCACCTCG AGGAGTCATC GATGCGCATG CGGAGGGGTG CCTCATCAGC	1320
GGCCCTATCG GCGGTTGCCA TCATCACAAC GGCCGTACGG GCACGGACAC TCACGATGTC	1380
TGACTCATCC CCCCACCTCG AGGAGTCATC GATGCGCATG CGGAGGGGTG CCTCATGAGC	1440
GCGGGCGGAC CCGCCTGGGT CATCACCGAT GACGCCCTCG CCCGCGAGGT GCTGGCCGAT	1500
CCCCGGTTCG TGAAGGACCC CGATCTCGCG CCCACCGCCT GCGGGGGGTG GGACGACGGT	1560
CTCGACATCC CCGTTCCGGA GCTGCGTCCG TTCACGCTCA TCGCCGTGGA CGGTGAGGAC	1620
CACCGCCGTC TGCGCCGCAT CCACGCACCG GCGTTCAACC CGCGCCGGCT GGCCGAGCGG	1680
ACGGATCGCA TCGCCGCCAT CGCCGACCGG CTGCTCACCG AACTCGCCGA CTCCTCCGAC	1740
CGGTCGGGCG AACCGGCCGA GCTGATCGGC GGCTTCGCGT ACCACTTCCC GCTGTTGGTC	1800
ATCTGCGAAC TGCTCGGCGT GCCGGTCACC GATCCGGCAA TGGCCCGCGA GGCCGTCCGC	1860
GTGCTCAAGG CACTCGGCCT CGGCGGCCCG CAGAGCGCCG GCGGTGACGG CACGGACCCT	1920
GCCGGGGACG TGCCGGACAC GTCGGCGCTG GAGAGCCTTC TCCTCGAAGC CGTGACGCG	1980

(2) INFORMATION FOR SEQ ID NO:7:

(A) LENGTH: 2081 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) **FEATURE:**

- (A) NAME/KEY: CDS
(B) LOCATION: 227..1649

TGTTGACAAT	TAATCATCCG	GCTCGTATAA	TGTGTGGAAT	TGTGAGCGGA	TAACAATTTC	60
ACACAGGAAA	CAGCGCCGCT	GAGAAAAAGC	GAAGCGGCAC	TGCTCTTTAA	CAATTTATCA	120
GACAATCTGT	GTGGGCCTC	GACCGGAATT	GGGCATCGAT	TAACTTTATT	ATTAAAAATT	180
AAAGAGGTAT	ATATTAATGT	ATCGATTAAA	TAAGGAGGAA	TAAACC	ATG GGG GGT	235
				Met	Gly Gly	
				1		
TCT CAT CAT CAT CAT CAT CAT	GGT ATG GCT AGC ATG ACT GGT GGA CAG	283				
Ser His His His His His His	Gly Met Ala Ser Met Thr Gly Gly Gln					
5	10	15				

-37-

CAA ATG GGT CGG GAT CTG TAC GAC GAT GAC GAT AAG GAT CGA TGG ATC Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Lys Asp Arg Trp Ile 20 25 30 35	331
CGA CCT CGA GAT CTG CAG ATG GTA CCA TAT GGG AAT TCG GAG GGG TGC Arg Pro Arg Asp Leu Gln Met Val Pro Tyr Gly Asn Ser Glu Gly Cys 40 45 50	379
CTC ATG AGC GGC GAG GCG CCG CGG GTG GCC GTC GAC CCG TTC TCG TGT Leu Met Ser Gly Glu Ala Pro Arg Val Ala Val Asp Pro Phe Ser Cys 55 60 65	427
CCC ATG ATG ACC ATG CAG CGC AAA CCC GAG GTG CAC GAC GCA TTC CGA Pro Met Met Thr Met Gln Arg Lys Pro Glu Val His Asp Ala Phe Arg 70 75 80	475
GAG GCG GGC CCC GTC GTC GAG GTG AAC GCC CCC GCG GGC GGA CCC GCC Glu Ala Gly Pro Val Val Glu Val Asn Ala Pro Ala Gly Gly Pro Ala 85 90 95	523
TGG GTC ATC ACC GAT GAC GCC CTC GCC CGC GAG GTG CTG GCC GAT CCC Trp Val Ile Thr Asp Asp Ala Leu Ala Arg Glu Val Leu Ala Asp Pro 100 105 110 115	571
CGG TTC GTG AAG GGA CCC GAT CTC GCG CCC ACC GCC TGG CGG GGG GTG Arg Phe Val Lys Gly Pro Asp Leu Ala Pro Thr Ala Trp Arg Gly Val 120 125 130	619
GAC GAC GGT CTC GAC ATC CCC GTT CCG GAG CTG CGT CCG TTC ACG CTC Asp Asp Gly Leu Asp Ile Pro Val Pro Glu Leu Arg Pro Phe Thr Leu 135 140 145	667
ATC GCC GTG GAC GGT GAG GAC CAC CGG CGT CTG CGC CGC ATC CAC GCA Ile Ala Val Asp Gly Glu Asp His Arg Arg Leu Arg Arg Ile His Ala 150 155 160	715
CCG GCG TTC AAC CCG CGC CGG CTG GCC GAG CGG ACG GAT CGC ATC GCC Pro Ala Phe Asn Pro Arg Arg Leu Ala Glu Arg Thr Asp Arg Ile Ala 165 170 175	763
GCC ATC GCC GAC CGG CTG CTC ACC GAA CTC GCC GAC TCC TCC GAC CGG Ala Ile Ala Asp Arg Leu Leu Thr Glu Leu Ala Asp Ser Ser Asp Arg 180 185 190 195	811
TCG GGC GAA CCG GCC GAG CTG ATC GGC GGC TTC GCG TAC CAC TTC CCG Ser Gly Glu Pro Ala Glu Leu Ile Gly Gly Phe Ala Tyr His Phe Pro 200 205 210	859
CTG TTG GTC ATC TGC GAA CTG CTC GGC GTG CCG GTC ACC GAT CCG GCA Leu Leu Val Ile Cys Glu Leu Leu Gly Val Pro Val Thr Asp Pro Ala 215 220 225	907
ATG GCC CGC GAG GCC GTC GGC GTG CTC AAG GCA CTC GGC CTC GGC GGC Met Ala Arg Glu Ala Val Gly Val Leu Lys Ala Leu Gly Leu Gly Gly 230 235 240	955
CCG CAG AGC GCC GGC GGT GAC GGC ACG GAC CCT GCC GGG GAC GTG CCG Pro Gln Ser Ala Gly Gly Asp Gly Thr Asp Pro Ala Gly Asp Val Pro 245 250 255	1003
GAC ACG TCG GCG CTG GAG AGC CTT CTC CTC GAA GCC GTG CAC GCG GCC Asp Thr Ser Ala Leu Glu Ser Leu Leu Leu Glu Ala Val His Ala Ala 260 265 270 275	1051
CGG CGG AAA GAC ACC CGG ACC ATG ACC CGC GTG CTC TAT GAA CGC GCA Arg Arg Lys Asp Thr Arg Thr Met Thr Arg Val Leu Tyr Glu Arg Ala 280 285 290 295	1099

-38-

280	285	290	
CAG GCA GAG TTC GGC TCG GTC TCC GAC GAC CAG CTC GTC TAC ATG ATC Gln Ala Glu Phe Gly Ser Val Ser Asp Asp Gln Leu Val Tyr Met Ile 295 300 305			1147
ACC GGA CTC ATC TTC GCC GGC CAC GAC ACC ACC GGC TCG TTC CTG GGC Thr Gly Leu Ile Phe Ala Gly His Asp Thr Thr Gly Ser Phe Leu Gly 310 315 320			1195
TTC CTG CTT GCG GAG GTC CTG GCG GGC CGT CTC GCG GCG GAC GCC GAC Phe Leu Leu Ala Glu Val Leu Ala Gly Arg Leu Ala Ala Asp Ala Asp 325 330 335			1243
GGG GAC GCC ATC TCC CGG TTC GTG GAG GAG GCG CTG CGC CAC CAC CCG Gly Asp Ala Ile Ser Arg Phe Val Glu Glu Ala Leu Arg His His Pro 340 345 350 355			1291
CCG GTG CCC TAC TCG TTG TGG AGG TTC GCT GCC ACG GAG GTG GTC ATC Pro Val Pro Tyr Ser Leu Trp Arg Phe Ala Ala Thr Glu Val Val Ile 360 365 370			1339
CGC GGT GTC CGG CTG CCC CGC GGA GCG CCG GTA CTG GTG GAC ATC GAG Arg Gly Val Arg Leu Pro Arg Gly Ala Pro Val Leu Val Asp Ile Glu 375 380 385			1387
GGC ACC AAC ACC GAC GGC CGC CAT CAC GAC GCC CCG CAC GCT TTC CAC Gly Thr Asn Thr Asp Gly Arg His His Asp Ala Pro His Ala Phe His 390 395 400			1435
CCG GAC CGC CCT TCG AGG CGG CGG CTC ACC TTC GGC GAC GGG CCG CAC Pro Asp Arg Pro Ser Arg Arg Arg Leu Thr Phe Gly Asp Gly Pro His 405 410 415			1483
TAC TGC ATC GGG GAG CAG CTC GCC CAG CTG GAA TCG CGC ACG ATG ATC Tyr Cys Ile Gly Glu Gln Leu Ala Gln Leu Glu Ser Arg Thr Met Ile 420 425 430 435			1531
GGC GTA CTG CGC AGC AGG TTC CCC CAA GCC CGA CTG GCC GTG CCG TAC Gly Val Leu Arg Ser Arg Phe Pro Gln Ala Arg Leu Ala Val Pro Tyr 440 445 450			1579
GAG GAG TTG CGG TGG TGC AGG AAG GGG GCC CAG ACA GCG CGG CTC ACT Glu Glu Leu Arg Trp Cys Arg Lys Gly Ala Gln Thr Ala Arg Leu Thr 455 460 465			1627
GAC CTG CCC GTC TGG CTG CGT T GATGGGCCGA CCGCGACCCG GCACGGGACC Asp Leu Pro Val Trp Leu Arg 470			1679
GGCCACCGCC CATCGCGCGG TGGGCGGTCC CGTGCCGGTC GCCCGGTGCG GTCCTCTCCC			1739
GACGCTCGCT CCCCCTGTGA CTTTCTCACA TCGAGACGTG ACGAAATAAT CCCAGCAAGT			1799
GCCATGCACA CTTTCATGGC GGACATTAC TTGCGAGGAT GGAGTGAGCA CACGGGGCCG			1859
CCCGAGACAC CCTACGGCCG CCGGAAGTAT GCCACCTGTT GACGCGAATG GAACGCCACA			1919
GAGGGAGCAC CGGCAATGCA GATCAATATG TTGGGCCCCG TCGTTGCACA TCACAATGGC			1979
ACGTCGGTGA CCCCAGATAGC CAGAAAACCC CGGCAGGTAT TCTCACTGCT CGCTCTTCAG			2039
GCAGGAACCG TCGTTCCGGT CCCCAGCTG ATGGAGGAGC TC			2081

(2) INFORMATION FOR SEQ ID NO:8:

-39-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 474 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Gly Gly Ser His His His His His His Gly Met Ala Ser Met Thr
 1 5 10 15
 Gly Gly Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Lys Asp
 20 25 30
 Arg Trp Ile Arg Pro Arg Asp Leu Gln Met Val Pro Tyr Gly Asn Ser
 35 40 45
 Glu Gly Cys Leu Met Ser Gly Glu Ala Pro Arg Val Ala Val Asp Pro
 50 55 60
 Phe Ser Cys Pro Met Met Thr Met Gln Arg Lys Pro Glu Val His Asp
 65 70 75 80
 Ala Phe Arg Glu Ala Gly Pro Val Val Glu Val Asn Ala Pro Ala Gly
 85 90 95
 Gly Pro Ala Trp Val Ile Thr Asp Asp Ala Leu Ala Arg Glu Val Leu
 100 105 110
 Ala Asp Pro Arg Phe Val Lys Gly Pro Asp Leu Ala Pro Thr Ala Trp
 115 120 125
 Arg Gly Val Asp Asp Gly Leu Asp Ile Pro Val Pro Glu Leu Arg Pro
 130 135 140
 Phe Thr Leu Ile Ala Val Asp Gly Glu Asp His Arg Arg Leu Arg Arg
 145 150 155 160
 Ile His Ala Pro Ala Phe Asn Pro Arg Arg Leu Ala Glu Arg Thr Asp
 165 170 175
 Arg Ile Ala Ala Ile Ala Asp Arg Leu Leu Thr Glu Leu Ala Asp Ser
 180 185 190
 Ser Asp Arg Ser Gly Glu Pro Ala Glu Leu Ile Gly Gly Phe Ala Tyr
 195 200 205
 His Phe Pro Leu Leu Val Ile Cys Glu Leu Leu Gly Val Pro Val Thr
 210 215 220
 Asp Pro Ala Met Ala Arg Glu Ala Val Gly Val Leu Lys Ala Leu Gly
 225 230 235 240
 Leu Gly Gly Pro Gln Ser Ala Gly Gly Asp Gly Thr Asp Pro Ala Gly
 245 250 255
 Asp Val Pro Asp Thr Ser Ala Leu Glu Ser Leu Leu Leu Glu Ala Val
 260 265 270
 His Ala Ala Arg Arg Lys Asp Thr Arg Thr Met Thr Arg Val Leu Tyr
 275 280 285
 Glu Arg Ala Gln Ala Glu Phe Gly Ser Val Ser Asp Asp Gln Leu Val
 290 295 300

-40-

Tyr Met Ile Thr Gly Leu Ile Phe Ala Gly His Asp Thr Thr Gly Ser
 305 310 315 320
 Phe Leu Gly Phe Leu Leu Ala Glu Val Leu Ala Gly Arg Leu Ala Ala
 325 330 335
 Asp Ala Asp Gly Asp Ala Ile Ser Arg Phe Val Glu Glu Ala Leu Arg
 340 345 350
 His His Pro Pro Val Pro Tyr Ser Leu Trp Arg Phe Ala Ala Thr Glu
 355 360 365
 Val Val Ile Arg Gly Val Arg Leu Pro Arg Gly Ala Pro Val Leu Val
 370 375 380
 Asp Ile Glu Gly Thr Asn Thr Asp Gly Arg His His Asp Ala Pro His
 385 390 395 400
 Ala Phe His Pro Asp Arg Pro Ser Arg Arg Arg Leu Thr Phe Gly Asp
 405 410 415
 Gly Pro His Tyr Cys Ile Gly Glu Gln Leu Ala Gln Leu Glu Ser Arg
 420 425 430
 Thr Met Ile Gly Val Leu Arg Ser Arg Phe Pro Gln Ala Arg Leu Ala
 435 440 445
 Val Pro Tyr Glu Glu Leu Arg Trp Cys Arg Lys Gly Ala Gln Thr Ala
 450 455 460
 Arg Leu Thr Asp Leu Pro Val Trp Leu Arg
 465 470

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 443 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Asp Arg Trp Ile Arg Pro Arg Asp Leu Gln Met Val Pro Tyr Gly Asn
 1 5 10 15
 Ser Glu Gly Cys Leu Met Ser Gly Glu Ala Pro Arg Val Ala Val Asp
 20 25 30
 Pro Phe Ser Cys Pro Met Met Thr Met Gln Arg Lys Pro Glu Val His
 35 40 45
 Asp Ala Phe Arg Glu Ala Gly Pro Val Val Glu Val Asn Ala Pro Ala
 50 55 60
 Gly Gly Pro Ala Trp Val Ile Thr Asp Asp Ala Leu Ala Arg Glu Val
 65 70 75 80
 Leu Ala Asp Pro Arg Phe Val Lys Asp Pro Asp Leu Ala Pro Thr Ala
 85 90 95
 Trp Arg Gly Val Asp Asp Gly Leu Asp Ile Pro Val Pro Glu Leu Arg
 100 105 110

-41-

Pro Phe Thr Leu Ile Ala Val Asp Gly Glu Asp His Arg Arg Leu Arg
 115 120 125
 Arg Ile His Ala Pro Ala Phe Asn Pro Arg Arg Leu Ala Glu Arg Thr
 130 135 140
 Asp Arg Ile Ala Ala Ile Ala Asp Arg Leu Leu Thr Glu Leu Ala Asp
 145 150 155 160
 Ser Ser Asp Arg Ser Gly Glu Pro Ala Glu Leu Ile Gly Gly Phe Ala
 165 170 175
 Tyr His Phe Pro Leu Leu Val Ile Cys Glu Leu Leu Gly Val Pro Val
 180 185 190
 Thr Asp Pro Ala Met Ala Arg Glu Ala Val Gly Val Leu Lys Ala Leu
 195 200 205
 Gly Leu Gly Gly Pro Gln Ser Ala Gly Gly Asp Gly Thr Asp Pro Ala
 210 215 220
 Gly Asp Val Pro Asp Thr Ser Ala Leu Glu Ser Leu Leu Leu Glu Ala
 225 230 235 240
 Val His Ala Ala Arg Arg Lys Asp Thr Arg Thr Met Thr Arg Val Leu
 245 250 255
 Tyr Glu Arg Ala Gln Ala Glu Phe Gly Ser Val Ser Asp Asp Gln Leu
 260 265 270
 Val Tyr Met Ile Thr Gly Leu Ile Phe Ala Gly His Asp Thr Thr Gly
 275 280 285
 Ser Phe Leu Gly Phe Leu Leu Ala Glu Val Leu Ala Gly Arg Leu Ala
 290 295 300
 Ala Asp Ala Asp Gly Asp Ala Ile Ser Arg Phe Val Glu Glu Ala Leu
 305 310 315 320
 Arg His His Pro Pro Val Pro Tyr Thr Leu Trp Arg Phe Ala Ala Thr
 325 330 335
 Glu Val Val Ile Arg Gly Val Arg Leu Pro Arg Gly Ala Pro Val Leu
 340 345 350
 Val Asp Ile Glu Gly Thr Asn Thr Asp Gly Arg His His Asp Ala Pro
 355 360 365
 His Ala Phe His Pro Asp Arg Pro Ser Arg Arg Arg Leu Thr Phe Gly
 370 375 380
 Asp Gly Pro His Tyr Cys Ile Gly Glu Gln Leu Ala Gln Leu Glu Ser
 385 390 395 400
 Arg Thr Met Ile Gly Val Leu Arg Ser Arg Phe Pro Gln Ala Arg Leu
 405 410 415
 Ala Val Pro Tyr Glu Glu Leu Arg Trp Cys Arg Lys Gly Ala Gln Thr
 420 425 430
 Ala Arg Leu Thr Asp Leu Pro Val Trp Leu Arg
 435 440

-42-

What is claimed is:

1. A method of making doxorubicin comprising the following steps:
 - a. providing a culture of a host microorganism transformed with a plasmid which contains *doxA* gene;
 - b. adding daunomycin to said cultures;
 - c. incubating said cultures in the presence of daunomycin; and extracting doxorubicin from said cultures.
2. The method of claim 1, wherein the host microorganism is bacterial.
3. The method of claim 1, wherein the host microorganism is *Streptomyces*.
4. The method of claim 1 wherein the *doxA* gene is isolated from *Streptomyces*.
5. The method of claim 1 wherein the *doxA* gene is driven by promoter selected from the group consisting of: *snpA*, *melCI*, or wild type promoter.
6. The method of claim 5, wherein the promoter is *snpA*.
7. The method of claim 6, wherein the promoter is activated by *Snpr*.
8. A gene which encodes daunomycin C-14 hydroxylase.
9. The gene of claim 8, wherein the gene encodes a protein having the amino acid sequence of Sequence Id. 5.
10. The gene of claim 8, wherein the gene is the *doxA* gene of *Streptomyces*.
11. The gene of claim 8, wherein the gene has the nucleotide sequence of Sequence ID. 4.
12. A genetically engineered host microorganism for converting daunomycin to doxorubicin comprising:
 - a. a plasmid, disposed within said microorganism comprising the following elements:
 - a *doxA* gene; and
 - a promoter driving the *doxA* gene;
 - b. a host microorganism.

-43-

13. A plasmid comprising the following elements:
- a *doxA* gene; and
 - a promoter driving the *doxA* gene.
14. The plasmid of claim 13, further comprising an activator for the promoter.
15. The plasmid of claim 13, further comprising a polylinker.
16. The plasmid of claim 15, wherein the polylinker has the following nucleotide sequence:
- GCATGCGAATTCAGATCTAGAGCTCAAGCTTTAACTAGTTAACGCGT SEQ.ID 3
17. A plasmid comprising the following elements:
- polylinker; and
 - a *snpA* promoter upstream of the polylinker.
18. A method of making doxorubicin comprising the following steps:
- providing a culture of *Streptomyces* sp. strain C5;
 - incubating said culture in the presence of daunomycin; and
 - extracting doxorubicin from said culture.
19. A method of making doxorubicin comprising the following steps:
- providing daunomycin C-14 hydroxylase;
 - incubating said daunomycin C-14 hydroxylase in the presence of daunomycin; and
 - extracting doxorubicin from the mixture of step b.
20. A method of making 13-deoxycarminomycin and 13-deoxydaunomycin comprising the following steps:
- providing a culture of a host microorganism transformed with a plasmid which contains *dauP* gene and *dauK*;
 - incubating said cultures in the presence of ϵ -Rhodomycin D; and
 - extracting 13-deoxycarminomycin and 13-deoxydaunomycin from said cultures.
21. An isolated daunomycin c-14 hydroxylase which converts of daunomycin to doxorubicin.
22. The plasmid of claim 17, wherein the polylinker has the following nucleotide sequence:
- GCATGCGAATTCAGATCTAGAGCTCAAGCTTTAACTAGTTAACGCGT SEQ.ID 3

-44-

23. The protein of claim 22, further comprising a leader sequence encoding six histidine residues,.

24. A method of producing 13-dihydrocarminomycin and carminomycin comprising the following steps:

- a. incubating daunomycin C-14 hydroxylase with 13-deoxycarminomycin; and
- b. extracting 13-dihydrocarminomycin and carminomycin.

25. A method of producing 13-dihydrodaunomycin and daunomycin comprising the following steps:

- a. incubating daunomycin C-14 hydroxylase with 13-deoxydaunomycin; and
- b. extracting 13-dihydrodaunomycin and daunomycin.

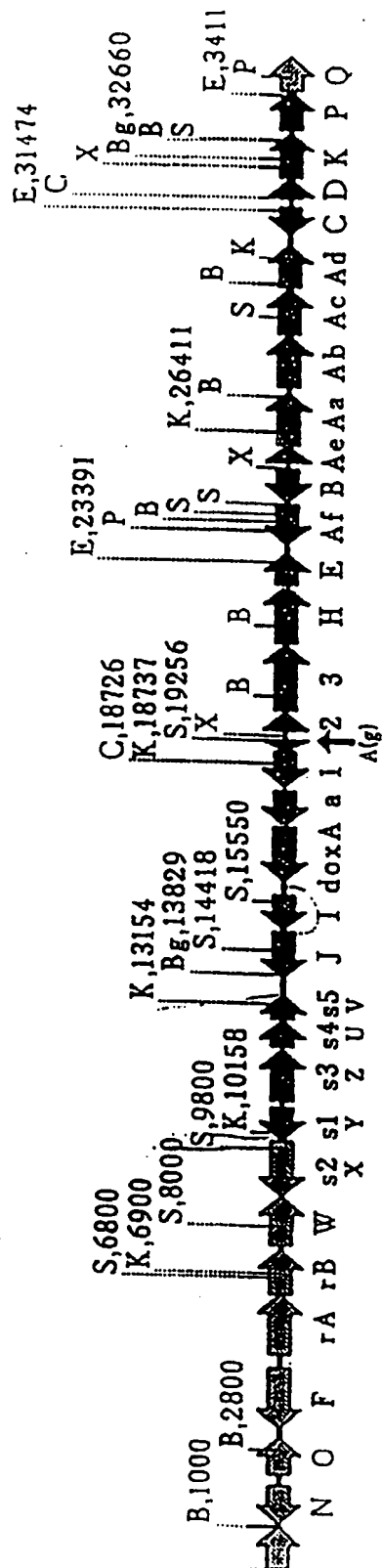


FIGURE 1

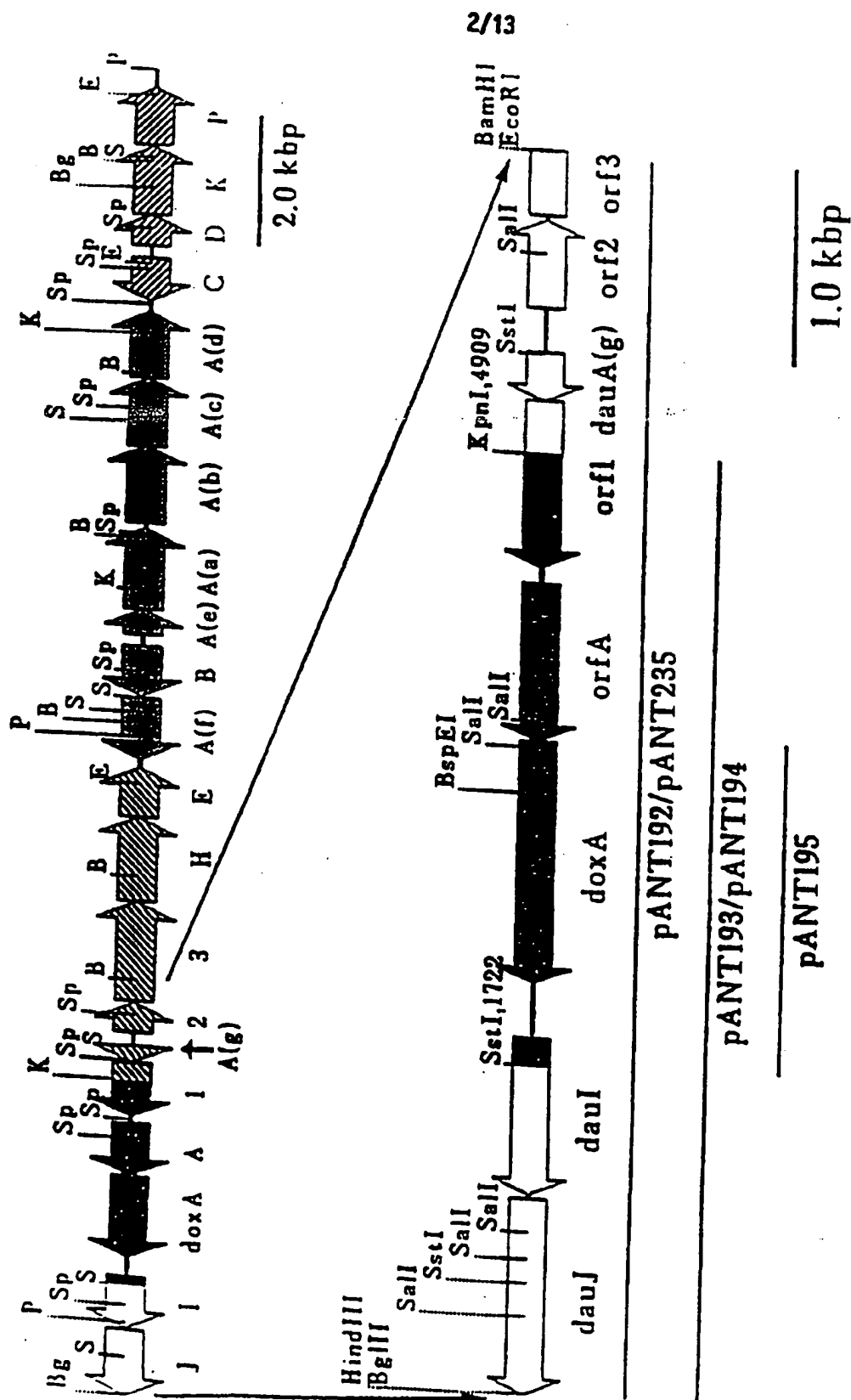


FIGURE 2

3/13

126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 3

[illegible]

FIGURE 3

FIGURE 4

5/13

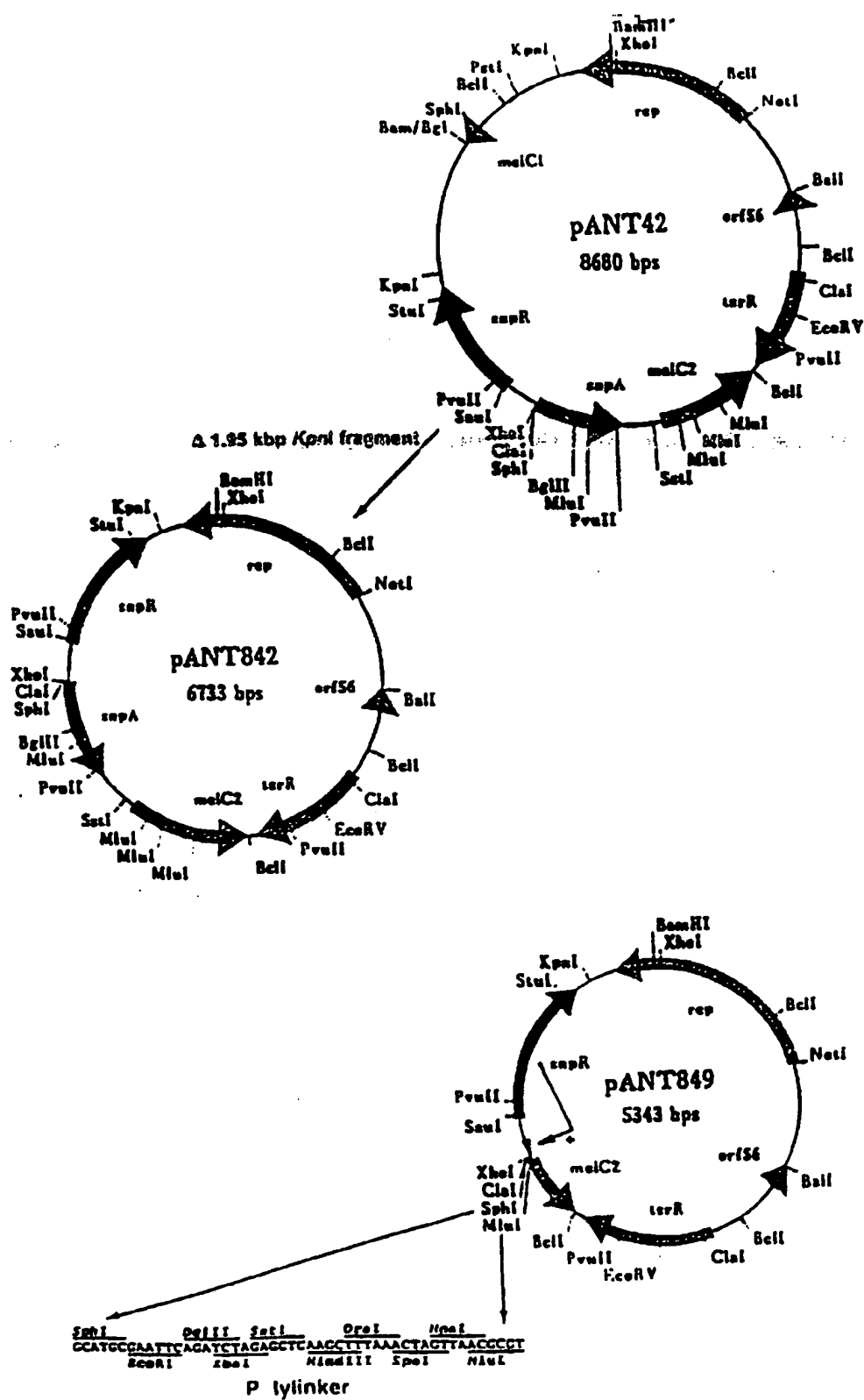
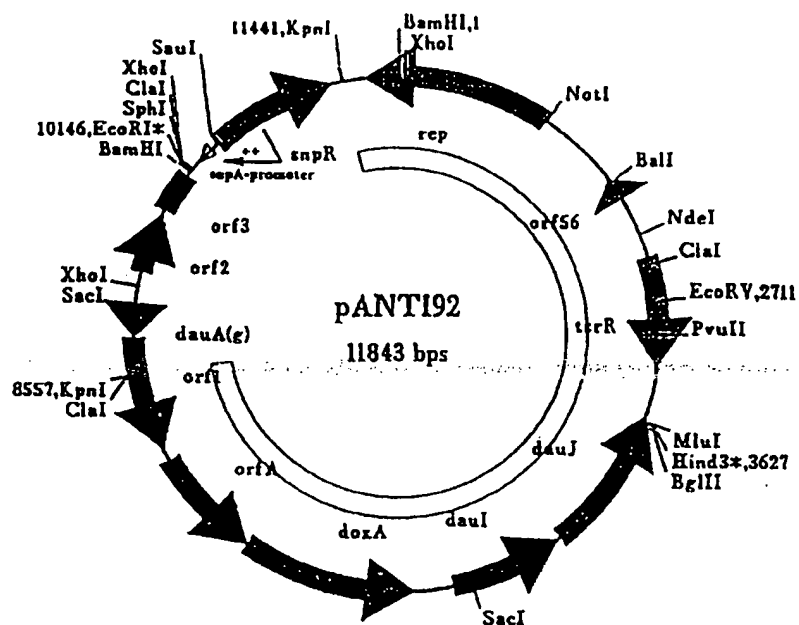
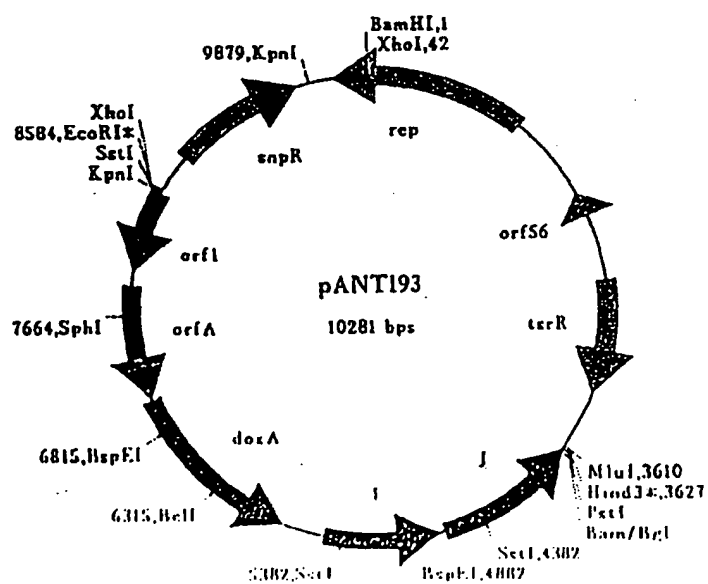


FIGURE 5

FIGURE 6

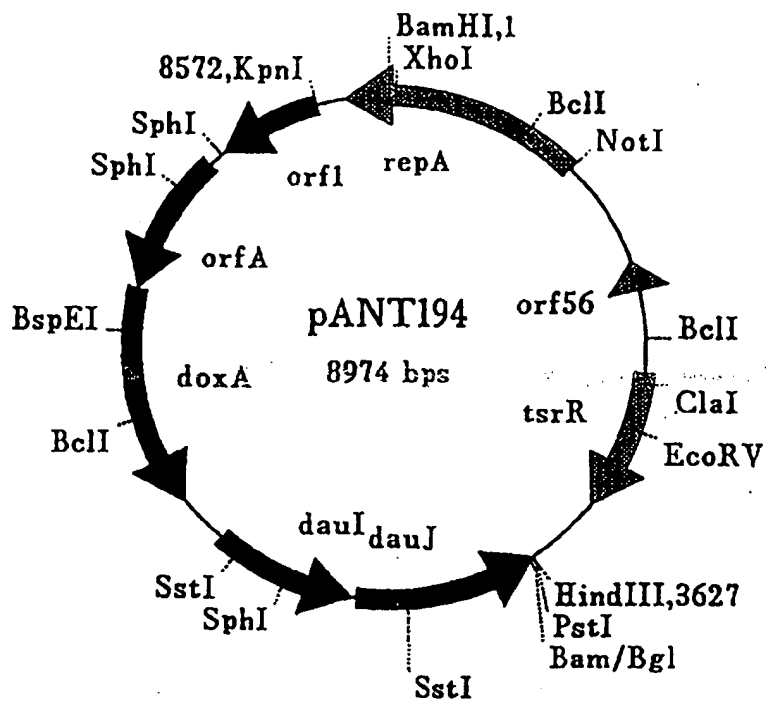
8/13
FIGURE 7

(Open box indicates sequence in pANT194)

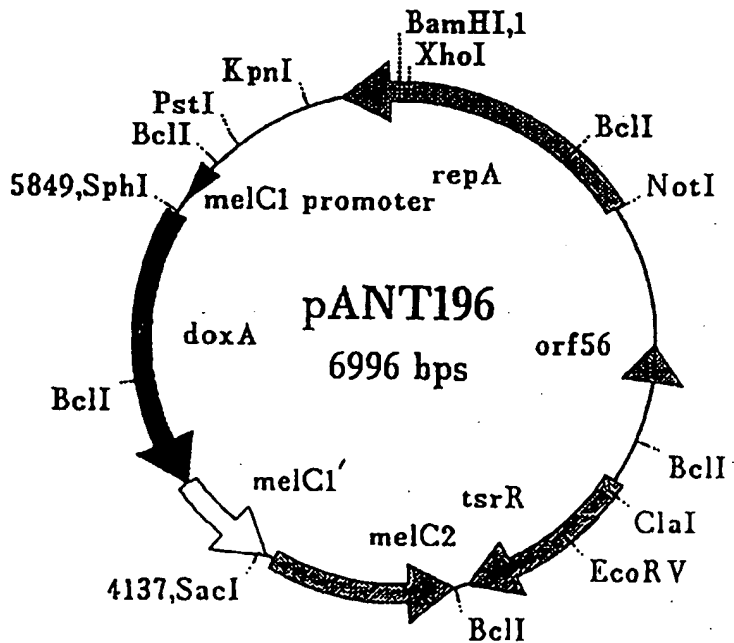


9/13

FIGURE 8



pANT196 - pIJ702 + *doxA* gene expressed from *melC1* promoter



```

-35          -10          lac operator
TGTGGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCGCCGCT      30

          rrnB antiterminator          g10 translational
GAGAAAAAGCGAAGCGGCACCTGCTCTTTAAACAATTATCAGACAATCTGTGTGGGCACTCGACCGGAATTGGGCATCGAT      160

enhancer          rbs          Mini cistron
TAACTTTATTATTAATAAATTAAAGAGGTATATATTAATGTATCGATTAAATAAGGAGGAATAAACCATGGGGGGTTCTCA      240
          FM Y R L N K E E *          FM G G S

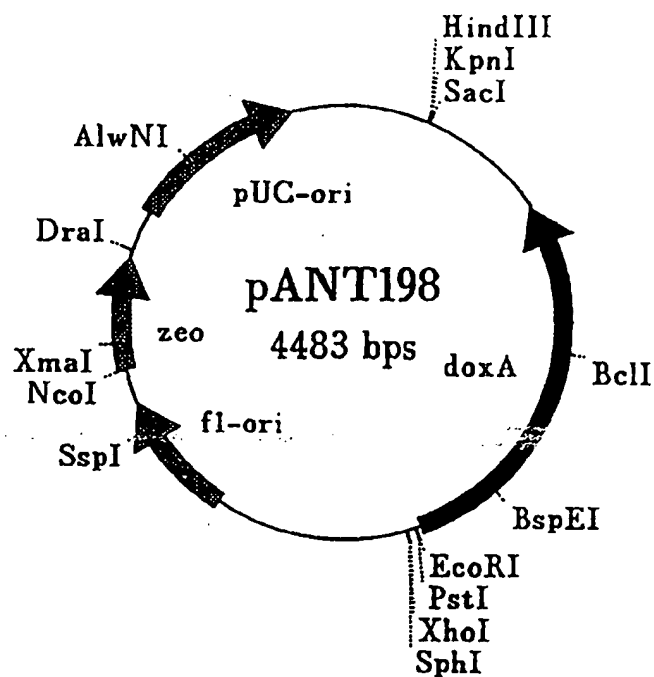
TCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAATGGGTCTGGGATCTGTACGACGATGACGATAAGG      320
H H H H H H H G M A S M T G G Q Q M G R D L Y D D D D K
          ↑
          Enterokinase cleavage site

          BglII
          BamHI          XhoI          PstI          KpnI          EcoRI
ATCGATGGATCCGACCTCGAGATCTGCAGATGGTACCATATGGGAATTCGGAGGGGTGCCTCATGACGGGCGAGGCGCCG...
D R W I R P R D L Q M V P Y G N S E G C L M S G E A P...
          Native DoxA sequence

```

FIGURE 9

11/13



pANT199 - doxA in pTrcHis for Expression
as a Fusion Protein in Escherichia coli

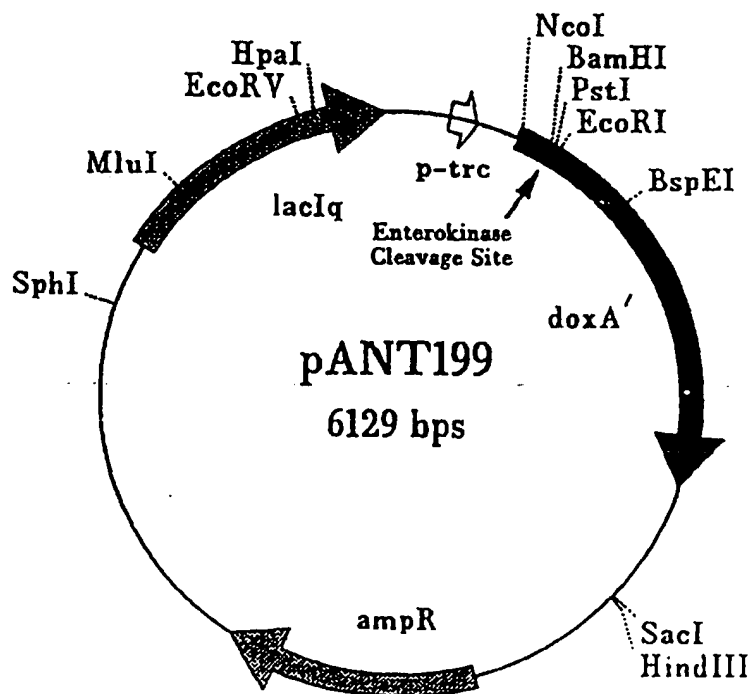


FIGURE 10

pANT144 - Used for bioconversion of
e-Rhodomyacin D to 13-Deoxycarminomycin
and 13-Deoxydaunomycin

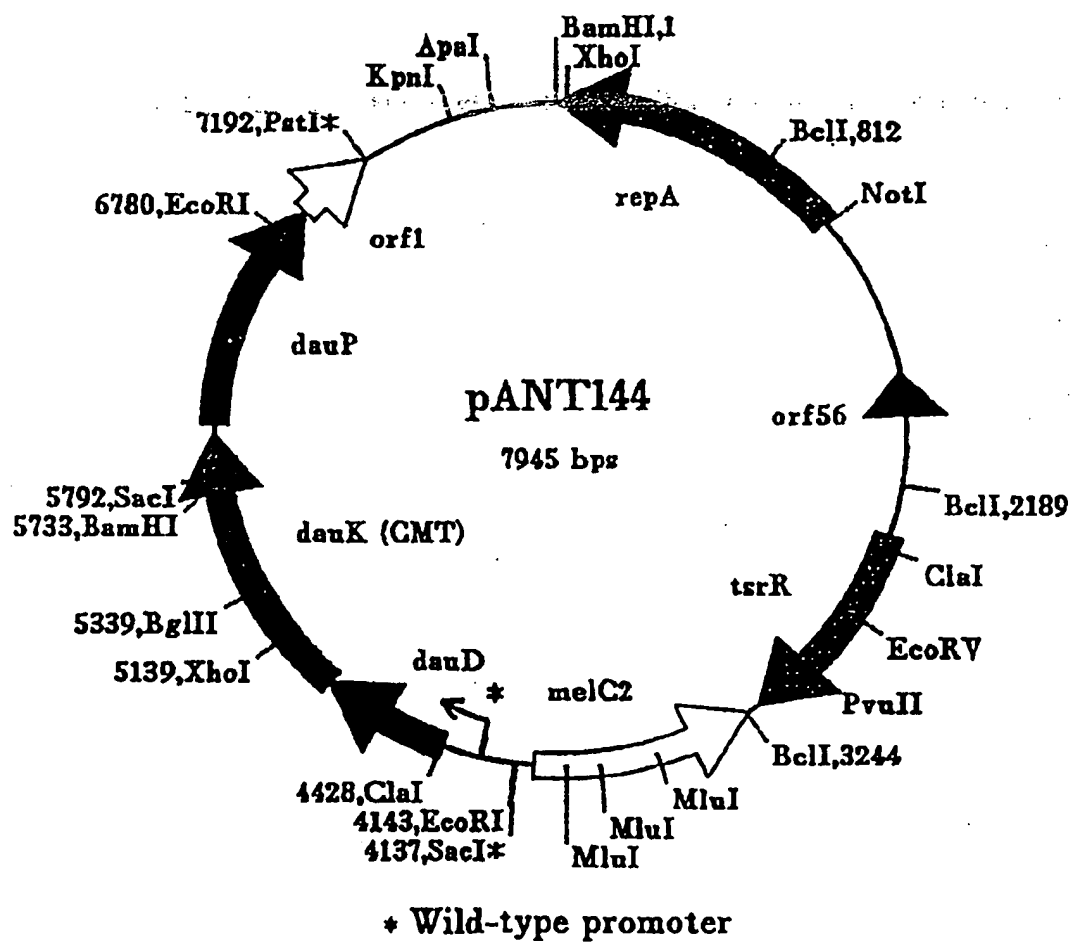


FIGURE 12